

# Permanental Point Processes on Real Tori, Theta Functions and Monge-Ampère Equations

Jakob Hultgren

May 19, 2016

## Abstract

Inspired by constructions in complex geometry we introduce a thermodynamic framework for Monge-Ampère equations on real tori. We show convergence in law of the associated point processes and explain connections to complex Monge-Ampère equations and optimal transport.

## Contents

<b>1</b>	<b>Introduction</b>	<b>1</b>
<b>2</b>	<b>Preliminaries: Optimal Transport on Real Tori</b>	<b>7</b>
<b>3</b>	<b>A Large Deviation Principle</b>	<b>15</b>
<b>4</b>	<b>The Rate Function and its relation to Monge Ampère equations</b>	<b>27</b>
<b>5</b>	<b>Existence and Uniqueness of Solutions</b>	<b>32</b>
<b>6</b>	<b>Geometric Motivation</b>	<b>37</b>

## 1 Introduction

In a series of papers culminating in [2] Berman introduce a thermodynamic framework for complex Monge-Ampère equations. In particular, he shows how the Monge-Ampère measures of solutions to complex Monge-Ampère equations can be seen as limits of canonically defined ( $\beta$ -deformations of) determinantal point processes. Inspired by this we will introduce a thermodynamic framework for real Monge-Ampère equations on the real torus  $X = \mathbb{R}^n / \mathbb{Z}^n$ . Using certain families of functions analogous to theta functions on Abelian varieties we construct permanental point processes on  $X$ . Our first result is that, as long as the Monge-Ampère equation admits a unique solution, the point processes defined by the statistic mechanical framework converges in law towards the Monge-Ampère measure of this unique solution. Equivalently, and in the language of

thermodynamics, under absence of first order phase transitions the microscopic setting admits a macroscopic limit that is determined by the Monge-Ampère equation.

The real torus should be seen as one of several settings where strong connections between complex geometry, real Monge-Ampère equations and optimal transport are manifested (the related case of toric manifolds is treated in [1]). We will exploit these connections to produce semi-explicit approximations of optimal transport maps on  $X$  (see Corollary 6). As such, this work ties in with the seminal works by McCann [22] and Cordero-Erasquin [11] on optimal transport on Riemannian manifolds.

Moreover, motivated by the difficult problem of singular Kähler-Einstein metrics of (almost everywhere) positive curvature on complex varieties we propose a corresponding real Monge-Ampère equation on  $X$  (see equation (3) below). The assumption of no first order phase transition always holds for positive temperature. However, a reflection of the fact that the related complex geometric problem is one of *positive* curvature is that the statistical mechanic setting for (3) is of *negative* temperature. As a second result, by proving a uniqueness theorem for Monge-Ampère equations of independent interest (see Theorem 2), we rule out first order phase transitions down to the critical temperature of  $-1$ . In a future paper we hope to address the question of uniqueness for temperatures smaller than  $-1$ , which might be seen as the analog of the problem studied in [18].

## 1.1 Setup

Let  $dx$  be the standard volume measure on  $X$  induced from  $\mathbb{R}^n$ . Let  $\beta$  be a real constant and  $\mu_0$  a probability measure on  $X$ , absolutely continuous and with smooth, strictly positive density with respect to  $dx$ . Given the data  $(\mu_0, \beta)$  we will consider the real Monge-Ampère equation on  $X$  given by

$$\text{MA}(\phi) = e^{\beta\phi} d\mu_0. \quad (1)$$

Here MA is the Monge-Ampère operator defined by

$$\phi \mapsto \det(\phi_{ij} + \delta_{ij}) dx. \quad (2)$$

where  $(\phi_{ij})$  is the Hessian of  $\phi$  with respect to the coordinates on  $X$  induced from  $\mathbb{R}^n$  and  $\delta_{ij}$  is the Kronecker delta. As usual we demand of a solution  $\phi : X \rightarrow \mathbb{R}$  that it is twice differentiable and quasi-convex in the sense that  $(\phi_{ij} + \delta_{ij})$  is a positive definite matrix.

We will pay specific attention to the case when  $\mu_0$  is chosen as the measure

$$\gamma = \sum_{m \in \mathbb{Z}^n} e^{-|x-m|^2/2} dx.$$

We get the equation

$$\text{MA}(\phi) = e^{\beta\phi} \gamma. \quad (3)$$

As mentioned above this equation has an interpretation in terms of complex geometry. For  $\beta = -1$ , (3) arises as the "push forward" of a twisted Kähler-Einstein equation on the Abelian variety  $\mathbb{C}^n/4\pi\mathbb{Z}^n + i\mathbb{Z}^n$ . A more detailed exposition of this relation will follow in Section 6.1.

## 1.2 Construction of the Point Processes

The point processes we will study arise as the so called " $\beta$ -deformations" of certain *permanental point processes* (see [19] for a survey). Let's first recall the general setup of a permanental point process with  $N$  particles. We begin by fixing a set of  $N$  *wave functions* on  $X$

$$S^{(N)} = \{\Psi_1^{(N)}, \dots, \Psi_N^{(N)}\}.$$

This defines a matrix valued function on  $X^N$

$$(x_1, \dots, x_N) \rightarrow (\Psi_i(x_j)).$$

Recall that the permanent of a matrix  $A = a_{ij}$  is the quantity

$$\sum_{\sigma} \prod_i a_{i, \sigma(i)}$$

where the sum is taken over all permutations of the set  $\{1 \dots, N\}$ . Together with the background measure  $\mu_0$  this defines a symmetric probability measure on  $X^N$

$$\text{perm}(\Psi_i^{(N)}(x_j)) d\mu_0^{\otimes N} / Z_N, \quad (4)$$

where  $Z_N$  is a constant ensuring the total mass is one. This is a *pure* permanental point process. We will define, for each  $k \in \mathbb{N}$ , a set of  $N = N_k$  wave functions and, for a given  $\beta \in \mathbb{R}$ , study the so called  $\beta$ -deformations of (4)

$$\mu_{\beta}^{(N)} = (\text{perm}(\Psi_i(x_j)))^{\beta/k} d\mu_0^{\otimes N} / Z_{\beta, N} \quad (5)$$

where, as above,  $Z_{\beta, N}$  is a constant ensuring the total mass is one. We will now define the sets of wave functions. Note that  $\mu_{\beta}^{(N)}$  does not depend on the order of the element in  $S^{(N)}$ . For each positive integer  $k$ , let

$$S^{(N)} = \{\Psi_p^{(N)} : p \in \frac{1}{k}\mathbb{Z}^n / \mathbb{Z}^n\}$$

where

$$\Psi_p^{(N)}(x) = \sum_{m \in \mathbb{Z}^n + p} e^{-k|x-m|^2/2} dx.$$

Before we move on we should make a comment on the notation. We get  $N = N_k = k^n$ . Throughout the text, in formulas where both  $N$  and  $k$  occur, the relation  $N = k^n$  will always be assumed.

Finally, we will make two remarks on the definitions. In [23] permanent point processes are used to model a bosonian many particle system in quantum mechanics. In that interpretation  $\Psi_i^{(N)}$  defines a 1-particle wave function and the permanent above is the corresponding  $N$ -particle wave function defined by  $\Psi_1^{(N)}, \dots, \Psi_N^{(N)}$ . Secondly, we will explain in Section 6.2 how the wave functions arises as the "push forward" of  $\theta$ -functions on  $\mathbb{C}^n/(4\pi\mathbb{Z}^n + i\mathbb{Z}^n)$ .

### 1.3 Main Results

Denote the space of probability measures on  $X$  by  $\mathcal{M}_1(X)$  and consider the map  $\delta^{(N)} : X^N \rightarrow \mathcal{M}_1(X)$

$$\delta^{(N)}(x) = \delta^{(N)}(x_1, \dots, x_N) = \frac{1}{N} \sum_{i=1}^N \delta_{x_i}.$$

Let  $x = (x_1, \dots, x_N) \in X^N$  be the random variable with law  $\mu_\beta^{(N)}$ . Its image under  $\delta^{(N)}$ ,  $\delta^{(N)}(x)$ , is the *empirical measure*. This is a random measure with law given by the push-forward measure

$$\Gamma_\beta^{(N)} = \left( \delta^{(N)} \right)_* \mu_\beta^{(N)} \in \mathcal{M}_1(\mathcal{M}_1(X)) \quad (6)$$

Our results concern the weak\* limit of  $\Gamma_\beta^{(N)}$  as  $N \rightarrow \infty$ . In particular we will show, in some cases, that the limit is a dirac measure concentrated at a certain  $\mu_* \in \mathcal{M}_1(X)$  related to (1) or (3). Loosely speaking, this means  $\mu_*$  can be approximated by sampling larger and larger point sets on  $X$  according to  $\mu_\beta^{(N)}$ .

**Theorem 1.** *Let  $\mu_0 \in \mathcal{M}_1(X)$  be absolutely continuous and have smooth, strictly positive density with respect to  $dx$ . Let  $\Gamma^{(N)}$  be defined as above and let  $\beta \in \mathbb{R}$ . Assume also that (1) admits a unique solution,  $\phi_*$ . Then*

$$\Gamma_\beta^{(N)} \rightarrow \delta_{\mu_*} \quad (7)$$

*in the weak\* topology of  $\mathcal{M}_1(\mathcal{M}_1(X))$ , where  $\mu_* = MA(\phi_*)$ .*

**Remark 1.** The assumption that (1) admits a unique solutions is always satisfied when  $\beta > 0$ . This follows from standard arguments (see Theorem 9). However, the case  $\beta < 0$  is a lot more subtle. In our second result we show that, in the special case  $\mu_0 = \gamma$ , the assumption holds for certain negative values of  $\beta$  as well.

**Theorem 2.** *Assume  $\mu_0 = \gamma$  and  $\beta \in [-1, 0)$ . Then equation (3) admits a unique solution.*

Note that if  $\beta \neq 0$  and  $\mu_* = MA(\phi_*)dx$  where  $\phi_*$  is a solution to (1), then  $\phi_*$  can be recovered from  $\mu_*$  as  $\phi_* = \frac{1}{\beta} \log \rho$  where  $\rho$  is the density of  $\mu_*$  with respect  $\mu_0$ . In fact we get the following corollary of Theorem 1.

**Corollary 1.** *Let  $\mu_0 \in \mathcal{M}_1(X)$  be absolutely continuous and have smooth, strictly positive density with respect to  $dx$ . Let  $\beta \neq 0$ . Assume also that (1) admits a unique solution,  $\phi_*$ . Let  $\phi_N : X \rightarrow \mathbb{R}$  be the function defined by*

$$\phi_N(x_1) = \frac{1}{\beta} \log \int_{X^{N-1}} \left( \text{perm}(\Psi_{p_i}^{(N)}(x_j)) \right)^{\beta/k} d\mu^{\otimes(N-1)}(x_2, \dots, x_n) / Z_{\beta, N}.$$

*Then  $\phi_N$  converges uniformly to  $\phi^*$ .*

If we put  $\beta = 0$  in (1) we get the inhomogenous Monge-Ampère equation. Solutions then determine Optimal Transport maps on  $X$ . Now, although Corollary 1 doesn't cover the case  $\beta = 0$ , by considering  $\mu_{\beta_N}^{(N)}$  for the sequence of constants  $\beta_N = 1/N$  we will be able to produce explicit approximations of optimal transport maps. However, when working with optimal transport it is natural to consider a more general setting than the one proposed for equation (1). Because of this we will not state this corollary here but postpone it to Section 6.3.

## 1.4 Outline

**Convergence in Theorem 1 and a Large Deviation Principle** Theorem 1 will follow from a *large deviation principle* for the sequence  $\Gamma^{(N)}$  (see Theorem 4). This large deviation principle provides a quantitative description of the convergence in Theorem 1, recording the speed of convergence in a *rate function*  $G : \mathcal{M}_1(X) \rightarrow [0, \infty)$ , satisfying  $\inf G = 0$  and a *rate*  $\{r_N\} \subset \mathbb{R}$  such that  $r_N \rightarrow \infty$  as  $N \rightarrow \infty$ . We will give a formal definition of large deviation principles in Section 3. Roughly speaking, a large deviation principle with rate function  $G$  and rate  $r_N$  holds if, for  $U \subset \mathcal{M}_1(X)$ , the probability  $\Gamma(U)$  behaves as

$$e^{-r_N \inf_U G}$$

as  $N \rightarrow \infty$ . This means  $\Gamma^{(N)}$ , for large  $N$ , is concentrated where  $G$  is small. In particular, if  $G$  admits a unique minimizer,  $\mu_*$ , (where  $G = 0$ ) then it follows that  $\Gamma^{(N)}$  converges in the weak\* topology to  $\delta_{\mu_*}$ .

**Proof of the Large Deviation Principle** It turns out that the rate function above is related to the Wasserstein metric of optimal transport. In Section 2 we will recall some basic facts about optimal transport. In particular, we explain how Kantorovich' duality principle gives an explicit formula for the Legendre transform of the squared Wasserstein distance from a fixed measure. The proof of Theorem 4 is given in Section 3 and it is divided into two parts of which the first part uses this explicit formula. In the first part, given in Section 3.1, we take a sequence of constants  $\beta_N$  such that  $\beta_N \rightarrow \infty$  and study the family  $\{\Gamma_{\beta_N}^{(N)}\}$ . In the thermodynamic interpretation this means we are studying the zero temperature limit of the system. Using the formula given by Kantorovich duality and the Gärtner-Ellis theorem, relating the moment generating functions of  $\Gamma_{\beta_N}^{(N)}$  to the Legendre transform of a rate function, we prove a large deviation

principle for this family (see Theorem 6). In the second part of the proof we show how the large deviation principle in Theorem 4 can be deduced from this. This is based on essentially well known arguments. However, for completeness we give a proof of this in Section 3.2. It turns out that the crucial point is the equicontinuity and uniform boundedness of the (normalized) energy functions

$$-\frac{1}{kN} \log \text{perm}(\Psi_i^{(N)}(x_j)).$$

These properties will follow from equicontinuity properties and bounds on the wave functions  $\Psi_i^{(N)}$  and we give a proof of these properties in Section 3.3.

**Connection to the Monge-Ampère Equation** The final ingredients in the proof of Theorem 1 are given in Section 4.1 and Section 4.2 (essentially by Lemma 11 and Theorem 13). These sections connect the large deviation principle above with the Monge-Ampère equation (1). Note that, as  $\inf G = 0$ ,  $G$  admits a unique point where  $G = 0$  if and only if  $G$  admits a unique minimizer. We apply a variational approach to (1). Uniqueness and existence of solutions is studied by means of a certain energy functional on  $C(X)$  whose stationary points corresponds to weak solutions of (1). The rate function above,  $G$ , is closely related to this energy functional. This relation encodes the fact that minimizers of  $G$  arise as the Monge-Ampère measures of solutions to (1). Moreover, it follows from this relation that  $G$  admits a unique minimizer if the energy functional does, which is true if and only if (1) admits a unique solution.

**Theorem 2** Existence of weak solutions will follow from the variational approach and compactness properties of the space of quasi convex functions on  $X$  (see Section 5.1) and regularity will follow from results by Caffarelli explained in Lemma 18. These type of existence results for Monge-Ampère equations on affine manifolds was originally proven by Caffarelli and Viaclovsky [9] on the one hand and Cheng and Yau [10] on the other. However, we will provide an alternative proof based on the variational principle above. Uniqueness, which is the main new contribution in this chapter is proved in Section 5.3. Here we look at the space of quasi-convex functions equipped with an affine structure different from the standard one. It will then follow from the Prekopa inequality that the energy functional associated to (3) is strictly convex with respect to this affine structure, hence admits no more than one minimizer. This is an extension of an argument used in [3] to prove uniqueness of Kähler-Einstein metrics on toric Fano manifolds. Curiously, there doesn't seem to be any direct argument for this using the Prekopa theorem on Riemannian manifolds (see [12]). Instead, we need to lift the problem to the covering space  $\mathbb{R}^n$  and use that  $\gamma$  is the push forward of a measure on  $\mathbb{R}^n$  with strong log-concavity properties.

**Geometric Motivation** In Section 6 we explain the connections to the point processes on compact Kähler manifolds introduced by Berman in [2]. More precisely, we explain the connection with a complex Monge-Ampère equations

on  $\mathbb{C}^n/4\pi\mathbb{Z}^n+i\mathbb{Z}^n$  and how the wave functions and permanental point processes defined here are connected to theta-functions and determinantal point processes on  $\mathbb{C}^n/4\pi\mathbb{Z}^n+i\mathbb{Z}^n$ . Finally, in Section 6.3 we show how the connection to optimal transport can be used to get explicit approximations of optimal transport maps on  $X$ .

We end this section with a comment. While some parts of Section 3 might be well known to readers with a probabilistic background and, likewise, some parts of Section 4 might be familiar to readers with a background in geometry or optimal transport we nevertheless want to provide a paper that is accessible to readers from all three of these fields. This should (at least partly) explain the length of the paper.

## 2 Preliminaries: Optimal Transport on Real Tori

In this section we will recall some basic theory of optimal transport. The content of the chapter is well known. Early contributors to the theory are Cordera-Erasquin [11] who established a theory of optimal transport on real tori and McCann [22] who took it to the very general setting of Riemannian manifolds. The reason for this is the close relation between optimal transport and real Monge-Ampère equations. The most important part is Corollary 2. There we explain how Kantorovich' duality theorem give a variational approach to real Monge-Ampère equations and an explicit formula for the Legendre transform of the functional  $\mu \rightarrow W^2(\mu, dx)$ , where  $W^2(\cdot, \cdot)$  is the Wasserstein metric, a distance function on  $\mathcal{M}_1(X)$  defined in terms of optimal transport and which turn up in the rate function describing the behaviour of the point process  $\Gamma^{(N)}$  as  $N \rightarrow \infty$ .

### 2.1 Kantorovich' Problem of Optimal Transport

We will use Kantorovich' formulation (as opposed to Monge's formulation) of the optimal transport problem. The given data is a smooth manifold  $Y$ , a *cost function*  $c : Y \times Y \rightarrow [0, \infty)$ , a *source measure*,  $\mu \in \mathcal{M}_1(Y)$  and a *target measure*,  $\nu \in \mathcal{M}_1(Y)$ . Kantorovich problem of optimal transport is the problem of minimizing the functional

$$C(\gamma) = \int_{Y \times Y} c(x, y) d\gamma(x, y)$$

over the set of *transport plans*,  $\Pi(\mu, dx)$ , consisting of measures  $\gamma \in \mathcal{M}_1(Y \times Y)$  such that the first and second marginals of  $\gamma$  equal  $\mu$  and  $\nu$  respectively. The optimal transport distance between  $\mu$  and  $\nu$  is the quantity

$$\inf_{\gamma \in \Pi(\mu, dx)} C(\gamma). \quad (8)$$

In our case  $Y = X$ ,  $\nu = dx$  and  $c = d(\cdot, \cdot)^2/2$  where  $d$  is the distance function on  $X$  induced from  $\mathbb{R}^n$ . In other words, if  $x, y \in \mathbb{R}^n$  and  $\pi : \mathbb{R}^n \rightarrow X$  is the

quotient map, then

$$c(\pi x, \pi y) = \frac{d(\pi x, \pi y)^2}{2} = \frac{\inf_{m \in \mathbb{Z}^n} |x - y - m|^2}{2}.$$

With this choice of cost function, (8) is often referred to as the (squared) Wasserstein distance,  $W^2(\mu, dx)$ , between  $\mu$  and  $dx$ .

## 2.2 The $c$ -Transform and $c$ -Convex Functions

A cost function in optimal transport defines a  $c$ -transform, closely related to Legendre transform on  $\mathbb{R}^n$ . Let  $C(X)$  be the space of continuous functions on  $X$ . For  $\phi \in C(X)$  the  $c$ -transform of  $\phi$  is

$$\phi^c(y) = \sup_{x \in X} -c(x, y) - \phi(x) = \sup_{x \in X} -\frac{d(x, y)^2}{2} - \phi(x) \quad (9)$$

Note that if  $\phi$  is a smooth function on  $X$  such that  $(\phi_{ij} + \delta_{ij})$  is positive definite, then there is a natural way of associating to  $\phi$  a convex function on  $\mathbb{R}^n$ , namely

$$\Phi(x) = \phi(\pi x) + \frac{x^2}{2}. \quad (10)$$

Let  $C(\mathbb{R}^n)$  be the space of continuous functions on  $\mathbb{R}^n$  and if  $\Phi \in C(X)$  let  $\Phi^*$  denote the Legendre transform of  $\Phi$ . The map from  $C(X)$  to  $C(\mathbb{R}^n)$  given by  $\phi \mapsto \Phi$ , relates  $c$ -transform on  $X$  to Legendre transform on  $\mathbb{R}^n$  in the sense that

**Lemma 1.** *Let  $\phi \in C(X)$  and*

$$\Phi(x) = \phi(\pi x) + \frac{x^2}{2}.$$

*Then*

$$\Phi^*(y) = \phi^c(\pi y) + \frac{y^2}{2}.$$

*Proof.* Note that

$$\begin{aligned} \sup_{x \in \mathbb{R}^n} -\frac{|x - y|^2}{2} - \phi(\pi x) &= \sup_{x \in [0, 1]^n, m \in \mathbb{Z}^n} -\frac{|x - y - m|^2}{2} - \phi(\pi x) \\ &= \sup_{x \in [0, 1]^n} - \inf_{m \in \mathbb{Z}^n} \frac{|x - y - m|^2}{2} - \phi(\pi x) \\ &= \sup_{x \in X} -\frac{d(x, \pi y)^2}{2} - \phi(x) \\ &= \phi^c(\pi y). \end{aligned}$$



This means

$$\begin{aligned}
\Phi^*(y) &= \sup_{x \in \mathbb{R}^n} \langle x, y \rangle - \Phi(x) \\
&= \sup_{x \in \mathbb{R}^n} -\frac{|x - y|^2}{2} - \phi(\pi x) + \frac{y^2}{2} \\
&= \phi^c(\pi y) + \frac{y^2}{2}.
\end{aligned}$$

which proves the lemma.  $\square$

It follows that  $\phi \in C(X)$  satisfies  $(\phi^c)^c = \phi$  if and only if  $\Phi$  is convex. The property  $(\phi^c)^c = \phi$  is often referred to as  $c$ -convexity and we will denote the set of functions in  $C(X)$  that satisfy this  $P(X)$ . Since  $\Phi^*$  is convex for any  $\Phi \in C(\mathbb{R}^n)$  we get that  $\phi^c \in P(X)$ , for any  $\phi \in C(X)$ . Moreover, also from the theory of convex functions on  $\mathbb{R}^n$ , we get that the projection  $\phi \mapsto (\phi^c)^c$  of  $C(X)$  onto  $P(X)$  is monotone in the sense that  $(\phi^c)^c(x) \leq \phi(x)$  for all  $x \in X$ .

Let  $P(\mathbb{R}^n)$  be the set of convex functions on  $\mathbb{R}^n$ . It is easy to verify that the image of  $P(X)$  in  $P(\mathbb{R}^n)$  under the map  $\phi \mapsto \Phi$  (where  $\Phi$  is given by (10)) is given by the set

$$\begin{aligned}
P_{\mathbb{Z}^n}(\mathbb{R}^n) &= \left\{ \Phi \in P(\mathbb{R}^n) : \Phi(x + m) - \frac{|x + m|^2}{2} = \Phi(x) - \frac{x^2}{2} \forall m \in \mathbb{Z}^n \right\} \\
&= \left\{ \Phi \in P(\mathbb{R}^n) : \Phi(x + m) = \Phi(x) + \langle x, m \rangle + \frac{m^2}{2} \forall m \in \mathbb{Z}^n \right\}
\end{aligned} \tag{11}$$

Now, let  $\phi \in P(X)$  and  $\Phi$  be the image of  $\phi$  in  $P_{\mathbb{Z}^n}(\mathbb{R}^n)$ . Then  $\Phi$  is differentiable at a point  $x \in \mathbb{R}^n$  if and only if  $\phi$  is differentiable at  $\pi x$ . Since a convex function on  $\mathbb{R}^n$  is differentiable almost everywhere we get that any  $\phi \in P(X)$  is differentiable almost everywhere (with respect to  $dx$ ). Further, it follows from (11) that  $\Phi$  is differentiable at  $x$  and  $\nabla \Phi(x) = y$  if and only if  $\Phi$  is differentiable at  $x + m$  and  $\nabla \Phi(x + m) = y + m$ . This means the map  $\nabla \Phi : \mathbb{R}^n \rightarrow \mathbb{R}^n$ , where it is defined, factors through to a map  $X \rightarrow X$ . This map is the so called  $c$ -gradient map in optimal transport, denoted  $\nabla^c \phi$ . It satisfies the formula

$$\nabla^c \phi(\pi x) = \pi \nabla \Phi(x).$$

Further,  $\Phi$  is differentiable at  $x$  and  $\Phi(x) = y$  if and only if  $y$  is the unique point in  $\mathbb{R}^n$  such that

$$\Phi(x) + \Phi^*(y) = \langle x, y \rangle. \tag{12}$$

This holds if and only if

$$\phi(\pi x) + \phi^c(\pi y) = -\frac{d(\pi x, \pi y)^2}{2}. \tag{13}$$

We conclude that  $\phi$  is differentiable and  $\nabla^c \phi(\pi x) = \pi y$  if and only if  $\pi y$  is the unique point in  $X$  such that (13) holds. In fact, this is the usual definition of

the  $c$ -gradient and one of its strengths is that it becomes immediately apparent that if  $\phi$  is differentiable at  $x$  and  $\phi^c$  is differentiable at  $y = \nabla^c \phi(x)$ , then  $\nabla^c \phi^c(y) = x$ .

The definition of the Monge-Ampère operator in (2) makes sense for twice differentiable functions. We will now provide an extension of this operator to  $P(X)$ .

**Definition 1.** Let  $\phi \in P(X)$ . We define the Monge-Ampère measure,  $\text{MA}(\phi)$ , of  $\phi$  as

$$\text{MA}(\phi) = (\nabla^c \phi^c)_* dx.$$

Consequently, we refer to functions in  $P(X)$  satisfying

$$(\nabla^c \phi^c)_* dx = \mu$$

as weak solutions to

$$\text{MA}(\phi) = \mu. \quad (14)$$

Now, the following lemma will serve as a direct justification of Definition 1 and we will see in Theorem 3 that it fits nicely into the theory of optimal transport. Moreover, weak solutions to (1) in terms of Definition 1 is the natural analog of so called Alexandrov solutions to Monge-Ampère equations on  $\mathbb{R}^n$  (see Section 5.2). In fact, we will see in Lemma 17 that the map  $\phi \mapsto \Phi$  where  $\Phi$  is given by (10) gives a direct link between these two types of solutions.

**Lemma 2.** Assume  $\phi$  is smooth and  $(\phi_{ij} + \delta_{ij})$  is strictly positive definite. Then

$$\det(\phi_{ij} + \delta_{ij}) dx = (\nabla^c \phi^c)_* dx.$$

*Proof.* First of all, we claim that  $\nabla^c \phi^c : X \rightarrow X$  is one-to-one. To see this, assume that  $\nabla^c \phi^c(x_1) = \nabla^c \phi^c(x_2)$  for  $x_1, x_2 \in X$ . Let  $\tilde{x}_1, \tilde{x}_2 \in \mathbb{R}^n$  be lifts of  $x_1$  and  $x_2$  respectively and  $\Phi^*$  be the image of  $\phi^*$  in  $P_{\mathbb{Z}^n}(\mathbb{R}^n)$ . We get

$$\nabla \Phi^*(\tilde{x}_1) = \nabla \Phi^*(\tilde{x}_2) + m.$$

By (11) we get  $\nabla \Phi^*(\tilde{x}_1) = \nabla \Phi^*(\tilde{x}_2 + m)$ . But since  $\phi$ , and hence  $\Phi$ , is smooth  $\Phi^*$  must be strictly convex. This means  $\tilde{x}_1 = \tilde{x}_2 + m$  and  $x_1 = x_2$ , proving the claim.

The previous claim implies, since  $\pi \circ \nabla \Phi^* = \nabla^c \phi^c \circ \pi$ , that  $\pi$  maps  $\nabla \Phi^*([0, 1]^n)$  diffeomorphically to  $X$ . Further,

$$\det(\phi_{ij} + \delta_{ij}) \circ \pi = \det(\Phi_{ij}^*) = \frac{1}{\det(\Phi_{ij}^*)} \quad (15)$$

and the numerator of the right hand side of (15) is the Jacobian determinant of the map  $\nabla \Phi^* : \mathbb{R}^n \rightarrow \mathbb{R}^n$ . Let  $h \in C(X)$ . Then

$$\begin{aligned} \int_X h \det(\phi_{ij} + \delta_{ij}) dx &= \int_{\nabla \Phi^*([0, 1]^n)} \frac{h \circ \pi}{\det(\Phi_{ij}^*)} dx = \int_{[0, 1]^n} h \circ \pi \circ \nabla \Phi^* dx \\ &= \int_{[0, 1]^n} h \circ \nabla^c \phi^c \circ \pi dx = \int_X h \circ \nabla^c \phi^c dx. \end{aligned} \quad (16)$$

which proves the lemma.  $\square$

### 2.3 Kantorovich Duality

We now return to the problem of optimal transport. Although it has very satisfactory solutions providing existence and characterization of minimizers under great generality, we will only give part of that picture here. For us, the important feature of the problem of optimal transport is its dual formulation. Introducing the functional  $\xi$  on  $C(X)$  defined by

$$\xi(\phi) = \int_X \phi^c dx$$

we get a functional  $J$  on  $C(X)$

$$J(\phi) = - \int_X \phi d\mu - \xi(\phi).$$

This functional describes the dual formulation of the problem of optimal transport in the sense that  $W^2(\mu, dx)$  can be recovered as the supremum of  $J$  over  $C(X)$ . Moreover, the maximizers of  $J$  are weak solutions to a certain Monge-Ampère equation. This is recorded in the following theorem.

**Theorem 3** ([20],[21],[5]). *Let  $\mu \in \mathcal{M}_1(X)$  be absolutely continuous with respect to  $dx$ . Let  $c = d^2/2$  where  $d$  is the distance function on  $X$  induced from  $\mathbb{R}^n$ . Then*

$$W^2(\mu, dx) = \inf_{\gamma \in \Pi(\mu, dx)} I(\gamma) = \sup_{\phi \in C(X)} J(\phi). \quad (17)$$

and there is  $\phi_\mu \in P(X)$  such that

$$\sup_{\phi \in C(X)} J(\phi) = J(\phi_\mu). \quad (18)$$

Moreover,

$$\text{MA}(\phi_\mu) = \mu. \quad (19)$$

**Remark 2.** Equation 17 is called Kantorovich' duality [20] and property (19) is the Knott-Smith criterion which, in the context of Monge's problem of optimal transport, was discovered independently by Knott and Smith in 1984 [21] and by Brenier in 1987 [5].

*Proof of Theorem 3.* The theorem is essentially given by Theorem 5.10 in [27]. As  $X$  is a smooth manifold that can be endowed with a complete metric,  $X$  is indeed a Polish space. Further,  $d$  is continuous and bounded on  $X$ . Putting  $\gamma' = \mu \times dx$  gives

$$\inf_{\gamma \in \mathcal{P}(\mu, dx)} I(\gamma) \leq I(\gamma') < \infty$$

hence the assumptions in 5.10.i, 5.10.ii and 5.10.iii in [27] holds. In particular we get that (17) holds and that there is an optimal transport plan  $\gamma \in \Pi(\mu, dx)$  and  $\phi_\gamma \in P(X)$  such that  $\gamma$  is concentrated on the set

$$\{(x, y) \in X \times X : \phi_\gamma(x) + \phi_\gamma^c(y) = -c(x, y)\}. \quad (20)$$

Let  $\phi_\mu = \phi_\gamma$ . To see that (18) holds, note that, since the first and second marginals of  $\gamma$  are  $\mu$  and  $\nu$  respectively,

$$\begin{aligned} W^2(\mu, dx) &= \int_{X \times X} c\gamma = - \int_{X \times X} (\phi_\mu(x) + \phi_\mu^c(y)) \gamma \\ &= - \int_X \phi_\mu(x) d\mu - \int_X \phi_\mu^c(y) dx. \end{aligned}$$

To see that (19) holds note that  $\phi_\mu^c \in P(X)$  is differentiable almost everywhere with respect to  $dx$ . Let  $A \subset X$  be a measurable set and  $\text{dom } \nabla^c \phi_\mu^c \subset X$  be the set where  $\phi_\mu^c$  is differentiable. We have

$$\gamma(X \times \text{dom } \nabla^c \phi_\mu^c) = dx(\text{dom } \nabla^c \phi_\mu^c) = 1.$$

As  $\gamma$  is concentrated on (20) we get that  $\gamma$  is concentrated on the set

$$\{(x, y) : y \in \text{dom } \nabla^c \phi_\mu^c, x = \nabla^c \phi_\mu^c(y)\}.$$

This means

$$\begin{aligned} \int_{(\nabla^c \phi_\mu^c)^{-1}(A)} dx &= \int_{X \times (\nabla^c \phi_\mu^c)^{-1}(A)} d\gamma = \int_{A \times (\nabla^c \phi_\mu^c)^{-1}(A)} d\gamma \\ &= \int_{A \times X} d\gamma = \int_A d\mu, \end{aligned}$$

in other words  $(\nabla^c \phi_\mu^c)_* dx = \mu$ , which proves (19).  $\square$

## 2.4 The Variational Approach to Real Monge-Ampère Equations

We will now reformulate the statement of Theorem 3 in terms of the Legendre transform and Gateaux differentiability of the functional  $\xi$ . Recall that if  $A$  is a functional on  $C(X)$ , then the *Legendre* transform of  $A$  is a functional on the dual vector space of  $C(X)$ , the space of finite signed measures on  $X$ ,  $\mathcal{M}(X)$ . This functional is given by

$$B(\mu) = \sup_{\phi \in C(X)} \int_Y \phi d\mu - A(\phi).$$

Recall also that if  $A$  is convex, then  $A$  is Gateaux differentiable at a point  $\phi$  and has Gateaux differential  $\mu$  if  $\mu$  is the unique point in  $\mathcal{M}(X)$  such that

$$B(\mu) = \int_Y \phi d\mu - A(\phi).$$

A priory  $W^2(\cdot, dx)$  is defined on  $\mathcal{M}_1(X)$ . However, we may extend the definition to all of  $\mathcal{M}(X)$  by putting  $W(\mu, dx) = +\infty$  for any  $\mu \notin \mathcal{M}_1(X)$ . We begin with the following lemma

**Lemma 3.** *The functional  $\xi$  is convex on  $C(X)$ . Moreover, let  $\phi_0, \phi_1 \in C(X)$  and*

$$\phi_t = t\phi_1 + (1-t)\phi_0.$$

*Then, if  $\xi(\phi_t)$  is affine in  $t$ ,*

$$\nabla^c \phi_0^c = \nabla^c \phi_1^c$$

*almost everywhere with respect to  $dx$ .*

*Proof.* First of all, for any  $y \in X$ , the quantity

$$\phi_t^c(y) = \sup_{x \in X} -c(x, y) - \phi_t(x) \quad (21)$$

is a supremum of functions that are affine in  $t$ , hence it is convex in  $t$ . This implies  $\xi(\phi_t)$  is convex in  $t$ . Now, assume  $\xi(\phi_t)$  is affine in  $t$ . This implies (21) is affine in  $t$  for almost all  $y$ . Assume  $y$  is a point such that  $\nabla^c \phi_0^c(y)$ ,  $\nabla^c \phi_{1/2}^c(y)$  and  $\nabla^c \phi_1^c(y)$  are defined and (21) is affine. Let  $x_{1/2} = \nabla^c \phi_{1/2}^c(y)$ . This means

$$\phi_{1/2}^c(y) = -c(x_{1/2}, y) - \phi_{1/2}(x_{1/2}).$$

By construction

$$\phi_t^c(y) \geq -c(x_{1/2}, y) - \phi_t(x_{1/2})$$

for any  $t \in [0, 1]$ . As  $\phi_t^c$  and  $-c(x_{1/2}, y) - \phi_t(x_{1/2})$  are affine functions (in  $t$ ) that coincide in one point in the interior of their domains, this inequality implies that they coincide. This means  $\nabla^c \phi_0^c(y) = \nabla^c \phi_{1/2}^c(y) = \nabla^c \phi_1^c(y)$ . As  $\nabla^c \phi_0^c$ ,  $\nabla^c \phi_{1/2}^c$  and  $\nabla^c \phi_1^c$  are defined almost everywhere, this proves the lemma.  $\square$

This allow us to draw the following conclusions from Theorem 3

**Corollary 2.** *The functional on  $\mathcal{M}(X)$  defined by  $\mu \mapsto W^2(-\mu, dx)$  is the Legendre transform of  $\xi$ . Moreover, for any  $\mu \in \mathcal{M}_1(X)$  there is  $\phi$  such that*

$$W^2(\mu, dx) + \xi(\phi) = - \int_X \phi d\mu.$$

*Finally,  $\xi$  is Gateaux differentiable on  $C(X)$  and*

$$d\xi|_\phi = -\text{MA}(\phi). \quad (22)$$

*Proof.* The first statement is, as long as  $\mu \in \mathcal{M}_1(X)$ , a direct consequence of (17). If  $\mu \notin \mathcal{M}_1(X)$  then putting  $\phi_C = \phi + C$  for some  $\phi \in C(X)$  and  $C \in \mathbb{R}$  gives  $(\phi_C)^c = \phi^c - C$  and

$$- \int_X \phi_C d\mu - \xi(\phi_C) = - \int_X \phi d\mu - \xi(\phi) + C(1 - \mu(X)).$$

Letting  $C \rightarrow \infty$  if  $\mu(X) < 1$  and  $C \rightarrow -\infty$  if  $\mu(X) > 1$  gives

$$\sup_{\phi \in C(X)} \phi d\mu - \xi(\phi) = +\infty,$$

proving the first statement. The second statement is also a direct consequence of Theorem 3. We will now prove that  $\xi$  is Gateaux differentiable and that (22) holds. Let  $\phi \in C(X)$ . We claim that there is  $\mu \in \mathcal{M}(X)$  such that

$$\xi(\phi) + W^2(\mu, dx) = - \int \phi d\mu, \quad (23)$$

in other words  $\mu$  is a supporting hyperplane of  $\xi$  at  $\phi$ . To see this, note that since  $W^2(\cdot, dx)$  is the Legendre transform of  $\xi$  we get that  $W^2(\cdot, dx)$  is lower semi-continuous and

$$\xi(\phi) + W^2(\mu, dx) \geq - \int \phi d\mu \quad (24)$$

for all  $\mu \in \mathcal{M}(X)$ . By lemma 3,  $\xi$  is convex on  $C(X)$ . By the involutive property of Legendre transform

$$\xi(\phi) = \sup_{\mu \in \mathcal{M}(X)} - \int \phi d\mu - W^2(\mu, dx).$$

Let  $\{\mu_i\} \subset \mathcal{M}(X)$  be a sequence such that

$$- \int \phi d\mu_i - W^2(\mu_i, dx) \rightarrow \xi(\phi).$$

We may assume, since  $W^2(\mu_i, dx) = \infty$  if  $\mu_i \notin \mathcal{M}_1(X)$ , that  $\mu_i \in \mathcal{M}_1(X)$  for all  $i$ . Since  $\mathcal{M}_1(X)$  is compact we may take a subsequence  $\{\mu_{i_k}\}$  converging to some  $\mu \in \mathcal{M}_1(X)$ . By the lower semi-continuity of  $W^2(\cdot, dx)$  we get

$$- \int \phi d\mu - W^2(\mu, dx) \geq \liminf_{k \rightarrow \infty} - \int \phi d\mu_{i_k} - W^2(\mu_{i_k}, dx) = \xi(\phi).$$

which, together with (24), proves the claim. We will now prove that this implies

$$(\nabla^c \phi^c)_* dx = \mu. \quad (25)$$

As this relation determines  $\mu$  we get that  $\mu$  must be the unique supporting hyperplane at  $\phi$ . This implies  $\xi$  is Gateaux differentiable at  $\phi$  and  $d\xi_\phi = \mu$ , proving the second statement in the corollary.

Now, to see that (25) holds, note that (23) implies  $W^2(\mu, dx) < \infty$  and hence  $\mu \in \mathcal{M}_1(X)$ . By Theorem 3 there is a function  $\phi_\mu \in P(X)$  such that  $\text{MA}(\phi_\mu) = \mu$  and

$$W^2(\mu, dx) + \xi(\phi_\mu) = - \int \phi_\mu d\mu.$$

This means  $\mu$  is a supporting hyperplane of  $\xi$  both at  $\phi$  and at  $\phi_\mu$ . This implies  $\xi(t\phi + (1-t)\phi_\mu)$  is affine. By Lemma 3,  $\nabla^c \phi^c$  and  $\nabla^c \phi_\mu^c$  coincide almost everywhere with respect to  $dx$  and hence (25) holds.  $\square$

### 3 A Large Deviation Principle

This section is devoted to Theorem 4 which will be the key part in the proof of Theorem 1. Before we state Theorem 4 we will recall the definition of the relative entropy function.

**Definition 2.** Assume  $\mu, \mu_0 \in \mathcal{M}(X)$  and, if  $\mu$  is absolutely continuous with respect to  $\mu_0$ , let  $\mu/\mu_0$  denote the density of  $\mu$  with respect to  $\mu_0$ . The relative entropy of  $\mu$  with respect to  $\mu_0$  is

$$Ent_{\mu_0}(\mu) = \begin{cases} \int_X \mu \log \frac{\mu}{\mu_0} & \text{if } \mu \text{ is a probability measure and absolutely} \\ & \text{continuous with respect to } \mu_0 \\ +\infty & \text{otherwise,} \end{cases}$$

We recall the basic property that  $Ent_{\mu_0}(\mu) \geq 0$  with equality if and only if  $\mu = \mu_0$ .

**Theorem 4.** Let  $\mu_0 \in \mathcal{M}_1(X)$  be absolutely continuous and have positive density with respect to  $dx$ . Let  $\beta \in \mathbb{R}$ . Assume  $\Gamma_\beta^{(N)}$  is defined as in section 1.2. Then

$$\left\{ \Gamma_\beta^{(N)} \right\}$$

satisfy a Large Deviation Principle with rate  $r_N = N$  and rate function

$$G(\mu) = \beta W^2(\mu, dx) + Ent_{\mu_0}(\mu) + C_{\mu_0, \beta}$$

where  $W^2(\mu, dx)$  is the squared Wasserstein 2-distance between  $dx$  and  $\mu_0$  (defined in the previous section) and  $C_{\mu_0, \beta}$  is a constant ensuring  $\inf_{\mathcal{M}_1(X)} G = 0$ .

Before we move on we will recall the definition of a Large Deviation Principle.

**Definition 3.** Let  $\chi$  be a topological space,  $\{\Gamma_N\}$  a sequence of probability measures on  $\chi$ ,  $G$  a lower semi continuous function on  $\chi$  and  $r_N$  a sequence of numbers such that  $r_N \rightarrow \infty$ . Then  $\{\Gamma_N\}$  satisfies a large deviation principle with rate function  $G$  and rate  $r_N$  if, for all measurable  $E \subset \chi$ ,

$$-\inf_{E^\circ} G \leq \liminf_{N \rightarrow \infty} \frac{1}{r_N} \log \Gamma_N(E) \leq \limsup_{N \rightarrow \infty} \frac{1}{r_N} \log \Gamma_N(E) \leq -\inf_{\bar{E}} G$$

where  $E^\circ$  and  $\bar{E}$  are the interior and the closure of  $E$ .

In our case  $\chi = \mathcal{M}_1(X)$ . As we may endow  $\mathcal{M}_1(X)$  with the Wasserstein 1-metric, metricizing the topology of weak\* convergence on  $\chi$ , we may think of  $\mathcal{M}_1(X)$  as a metric space. Further, by Prohorov's Theorem,  $\mathcal{M}_1(X)$  is compact. In this setting there is an alternative, and well known, criteria for when a large deviation principle exist.

**Lemma 4.** Let  $\chi$  be a compact metric space,  $\{\Gamma_N\}$  a sequence of probability measures on  $\chi$ ,  $G$  a function on  $\chi$  and  $r_N$  a sequence of numbers such that

$r_N \rightarrow \infty$ . Let  $B_d(\mu)$  denote the open ball in  $\chi$  with center  $\mu$  and radius  $d$ . Then  $\{\Gamma_N\}$  satisfies a large deviation principle with rate function  $G$  and rate  $r_N$  if and only if, for all  $\mu \in \chi$

$$\begin{aligned} G(\mu) &= \lim_{\delta \rightarrow 0} \limsup_{N \rightarrow \infty} -\frac{1}{r_N} \log \Gamma_N(B_\delta(\mu)) \\ &= \lim_{\delta \rightarrow 0} \liminf_{N \rightarrow \infty} -\frac{1}{r_N} \log \Gamma_N(B_\delta(\mu)) \end{aligned}$$

*Proof.* Let  $\mathcal{B}$  be the basis of the topology on  $\chi$  given by

$$\mathcal{B} = \{B_d(\mu) : d > 0, \mu \in \chi\}.$$

By Theorem 4.1.11, Theorem 4.1.18 and Lemma 1.2.18 (recall that  $\chi$  is compact by assumption) in [13],  $\{\Gamma_N\}$  satisfies a large deviation principle with rate function  $G$  and rate  $r_N$  if and only if

$$\begin{aligned} G(\mu) &= \sup_{B \in \mathcal{B}: \mu \in B} \limsup_{N \rightarrow \infty} -\frac{1}{r_N} \log \Gamma_N(B_\delta(\mu)) \\ &= \sup_{B \in \mathcal{B}: \mu \in B} \liminf_{N \rightarrow \infty} -\frac{1}{r_N} \log \Gamma_N(B_\delta(\mu)). \end{aligned}$$

Now, if  $\mu \in B \in \mathcal{B}$  then  $B_d(\mu) \subset B$  for  $d$  small enough. This means, since

$$\lim_{d \rightarrow 0} \limsup_{N \rightarrow \infty} -\frac{1}{r_N} \log \Gamma_N(B_\delta(\mu)) \quad (26)$$

is increasing as  $d \rightarrow 0$ , that

$$(26) \geq \sup_{B \in \mathcal{B}: \mu \in B} \limsup_{N \rightarrow \infty} -\frac{1}{r_N} \log \Gamma_N(B_\delta(\mu)). \quad (27)$$

Since, for any  $d > 0$ ,  $B_d(\mu)$  is a candidate for the supremum in the right hand side of (27) we get that equality must hold in (27). The same argument goes through with  $\limsup$  replaced by  $\liminf$ . This proves the lemma.  $\square$

Finally we recall the well known

**Theorem 5** (Sanov's theorem, see for example 6.2.10 in [13]). *Let  $\mu_0 \in \mathcal{M}_1(X)$ . Then the family*

$$\left\{ \left( \delta^{(N)} \right)_* \mu_0^{\otimes N} \right\}$$

*satisfies a large deviation principle with rate  $r_N = N$  and rate function  $\text{Ent}_{\mu_0}$ .*

### 3.1 The Zero Temperature Case and the Gärtner-Theorem

Recall that  $N = k^n$ . For each  $\beta \in \mathbb{R}$  we get a family of probability measures  $\{\Gamma_\beta^{(N)}\}_{k \in \mathbb{N}}$ . Theorem 1 and Theorem 4 are both concerned with the behavior of these families. In this section we will consider the family  $\{\Gamma_k^{(N)}\}_{k \in \mathbb{N}}$ . We will prove a large deviation principle for this family (see Theorem 6) which, in Section 3.2, will be used to prove Theorem 4.



**Theorem 6.** *Let  $\mu_0 \in \mathcal{M}_1(X)$  be absolutely continuous and have positive density with respect to  $dx$ . Assume  $\Gamma_\beta^{(N)}$  is defined as in section 1.2. Then*

$$\{\Gamma_k^{(N)}\}$$

*satisfies a large deviation principle with rate  $r_N = kN$  and rate function*

$$G(\mu) = W^2(\cdot, dx).$$

Recall that if  $\Gamma$  is a probability measure on a topological vector space  $\chi$ , then the moment generating function of  $\Gamma$  is the functional on the dual vector space  $\chi^*$  given by

$$Z_\Gamma(\phi) = \int_\chi e^{-\langle \phi, \mu \rangle} d\Gamma(\mu)$$

where  $\langle \cdot, \cdot \rangle$  is the pairing of  $\chi$  and  $\chi^*$ . The significance of this for our purposes lies in the Gärtner-Ellis theorem. Before we state this theorem, recall that a sequence of (Borel) probability measures  $\{\Gamma_N\}$  on a space  $\chi$  is *exponentially tight* if for each  $\epsilon \in \mathbb{R}$  there is a compact  $K_\epsilon \subseteq \chi$  such that for all  $N$

$$\limsup_{N \rightarrow \infty} \frac{1}{N} \log \Gamma_N(\chi \setminus K_\epsilon) \leq \epsilon. \quad (28)$$

In our case, when  $\chi$  is compact, this is automatically satisfied since choosing  $K_\epsilon = \chi$  for any  $\epsilon$  gives that the left hand side of (28) is  $-\infty$  for all  $N$ .

**Theorem 7** (The Gärtner-Ellis Theorem. See for example Corollary 4.5.27 in [13]). *Let  $\chi$  be a locally convex topological vector space,  $\{\Gamma_N\}$  an exponentially tight sequence of probability measures on  $\chi$  and  $r_N$  a sequence such that  $r_N \rightarrow \infty$ . Let  $Z_{\Gamma_N}$  be the moment generating function of  $\Gamma_N$  and assume*

$$F(\phi) = \lim_{N \rightarrow \infty} \frac{1}{r_N} \log Z_{\Gamma_N}(r_N \phi)$$

*exist, is finite valued, lower semi continuous and Gateaux differentiable. Then  $\Gamma_N$  satisfies a large deviation principle with rate  $r_N$  and rate function given by the Legendre transform of  $F$ .*

Theorem 6 will follow from the Gärtner-Ellis theorem and the crucial point will be the following lemma.

**Lemma 5.** *Let  $\mu_0 \in \mathcal{M}_1(X)$  be absolutely continuous and have positive density with respect to  $dx$ . Assume  $\Gamma_\beta^{(N)}$  is defined as in section 1.2. Then*

$$\lim_{N \rightarrow \infty} \frac{1}{kN} \log Z_{\Gamma^{(N)}}(kN\phi) = \xi(-\phi).$$

*Proof.* Note that if  $\mu_N$  is a measure on  $X^N$  and  $F$  is a function on  $\mathcal{M}_1(X)$ , then, since  $\Gamma^{(N)} = (\delta^{(N)})_* \mu_{\beta_N}^{(N)}$ ,

$$\int_{\mathcal{M}_1(X)} F(\mu) \Gamma^{(N)} = \int_{X^N} F(\delta^{(N)}(x)) d\mu_{\beta_N}^{(N)}.$$

Moreover,

$$\left\langle kN\phi, \delta^{(N)}(x) \right\rangle = kN \int_X \phi \frac{1}{N} \sum \delta_{x_i} = k \sum \phi(x_i).$$

This means

$$Z_{\Gamma^{(N)}}(kN\phi) = \int_{\mathcal{M}_1(X)} e^{\langle r_N\phi, \mu \rangle} \Gamma^{(N)} = \int_{X^N} e^{k \sum \phi(x_i)} d\mu_{\beta_N}^{(N)}.$$

Using the symmetries in the explicit form of  $\mu_{\beta_N}^{(N)}$  we get

$$\begin{aligned} Z_{\Gamma^{(N)}}(kN\phi) &= \int_{X^N} \sum_{\sigma} \prod_i \Psi_{p_i}^{(N)}(x_{\sigma(i)}) e^{k\phi(x_{\sigma(i)})} d\mu_0^{\otimes N} \\ &= \sum_{\sigma} \int_{\sigma^{-1}(X^N)} \prod_i \Psi_{p_i}^{(N)}(x_i) e^{k\phi(x_i)} d\mu_0^{\otimes N} \\ &= N! \int_{X^N} \prod_i \Psi_{p_i}^{(N)}(x_i) e^{k\phi(x_i)} d\mu_0^{\otimes N} \\ &= N! \prod_i \int_X \Psi_{p_i}^{(N)}(x) e^{k\phi(x)} d\mu_0 \end{aligned} \quad (29)$$

Introducing the notation

$$c_p^{(N)} = -\frac{1}{k} \log \Psi_p^{(N)}$$

we get

$$Z_{\Gamma^{(N)}}(kN\phi) = N! \prod_{p \in \frac{1}{k}\mathbb{Z}^n / \mathbb{Z}^n} \int_X e^{k(-c_p^{(N)} + \phi)} d\mu_0. \quad (30)$$

Now, we claim that

$$c_p^{(N)} \rightarrow d(x, p)^2 / 2 \quad (31)$$

uniformly in  $p$  and  $x$ . To see this, note first that

$$d(x, p)^2 = \inf_{m \in \mathbb{Z}^n + p} |x - m|^2$$

and

$$\begin{aligned} c_p^{(N)}(x) &= -\frac{1}{k} \log \sum_{m \in \mathbb{Z}^n + p} e^{-k|x-m|^2/2} \\ &\leq -\frac{1}{k} \log \sup_{m \in \mathbb{Z}^n + p} e^{-k|x-m|^2/2} = \inf_{m \in \mathbb{Z}^n + p} |x - m|^2 / 2. \end{aligned}$$

On the other hand, by the exponential decay of  $e^{-|x-m|^2}$  there is a large constant,  $C$ , (independent of  $x$  and  $p$ ) such that

$$\sum_{m \in \mathbb{Z}^n + p} e^{-k|x-m|^2/2} \leq C \sup e^{-k|x-m|^2/2}$$

and

$$\begin{aligned} c_p^{(N)}(x) &= -\frac{1}{k} \log \sum_{m \in \mathbb{Z}^n + p} e^{-k|x-m|^2/2} \geq -\frac{1}{k} \log \left( C \sup_{m \in \mathbb{Z}^n + p} e^{-k|x-m|^2/2} \right) \\ &= -\frac{\log C}{k} + \inf_{m \in \mathbb{Z}^n + p} |x-m|^2/2. \end{aligned}$$

This proves the claim. We claim further that

$$\frac{1}{k} \log \int_X e^{k(-c_p^{(N)} + \phi)} d\mu_0 \rightarrow (-\phi)^c(p) \quad (32)$$

uniformly in  $p$ . To see this, note first that (31) together with the fact that the family  $\{d(\cdot, p)^2/2 : p \in X\}$  is equi-continuous implies that

$$\{c_p^{(N)} : k \in \mathbb{N}, p \in X\}$$

is equi-continuous. This means for any  $\epsilon > 0$  there is  $d > 0$  such that for all  $k \in \mathbb{N}$  and  $p, x_* \in X$

$$|c_p^{(N)}(x) - \phi(x) - (c_p^{(N)}(x_*) - \phi(x_*))| \leq \epsilon \quad (33)$$

as long as  $x \in B_d(x_*)$ . Further, as  $\mu_0$  has full support, is absolutely continuous and has smooth density with respect to  $dx$  there is a large constant  $C$  such that

$$C\mu_0(B_d(x_*)) \geq 1 \quad (34)$$

for all  $x_* \in X$ . We get trivially

$$\begin{aligned} \frac{1}{k} \log \int_X e^{k(-c_p^{(N)} + \phi)} d\mu_0 &\leq \frac{1}{k} \log \sup_{x \in X} e^{k(-c_p^{(N)} + \phi)} \\ &= \sup_{x \in X} -c_p^{(N)}(x) + \phi(x) \end{aligned} \quad (35)$$

For each  $N$ , let  $x_*^{(N)}$  satisfy

$$-c_p^{(N)}(x_*^{(N)}) + \phi(x_*^{(N)}) = \sup_{x \in X} -c_p^{(N)}(x) + \phi(x).$$

Using (33) and (34) gives

$$\begin{aligned} \frac{1}{k} \log \int_X e^{k(-c_p^{(N)} + \phi)} d\mu_0 &\geq \frac{1}{k} \log \int_{B_\delta(x_*^{(N)})} e^{k(\sup_{x \in X} -c_p^{(N)} + \phi - \epsilon)} d\mu_0 \\ &= \frac{1}{k} \log \int_{B_\delta(x_*^{(N)})} d\mu_0 + \sup_{x \in X} -c_p^{(N)}(x) + \phi(x) - \epsilon \\ &\geq \frac{1}{k} \log \frac{1}{C} \int_X d\mu_0 + \sup_{x \in X} -c_p^{(N)}(x) + \phi(x) - \epsilon. \end{aligned} \quad (36)$$

Finally, letting  $k, N \rightarrow \infty$  and  $\epsilon \rightarrow 0$  in (35) and (36) proves (32). Recalling equation (30), we have

$$\begin{aligned} \frac{1}{kN} \log Z_{\Gamma(N)}(kN\phi) &= \frac{1}{kN} \log N! \prod_{p \in \frac{1}{k}\mathbb{Z}^n/\mathbb{Z}^n} \int_X e^{k(-c_p^{(N)} + \phi)} d\mu_0 \\ &= \frac{\log N!}{kN} + \frac{1}{N} \sum_{p \in \frac{1}{k}\mathbb{Z}^n/\mathbb{Z}^n} \frac{1}{k} \log \int_X e^{k(-c_p^{(N)} + \phi)} d\mu_0 \end{aligned} \quad (37)$$

By Sterling's formula,  $\log N! \leq N \log N + O(\log N)$ . This means, since  $N = k^n$ , that the first term in (37) is bounded by  $(\log k^n)/k + O(\log k^n)/k^{n+1}$  which vanishes as  $k \rightarrow \infty$ . Finally, using (32) we get, since  $\frac{1}{N} \sum_{p \in \frac{1}{k}\mathbb{Z}^n/\mathbb{Z}^n} \delta_p \rightarrow dx$  in the weak\* topology, that the second term converges to

$$\int (-\phi)^c(p) dx = \xi(-\phi).$$

This proves the lemma.  $\square$

When proving Theorem 6 we will also need the following lemma.

**Lemma 6.** *The functional  $\xi$  is continuous on  $C(X)$ .*

*Proof.* We will prove that for any  $\phi_0, \phi_1 \in C(X)$

$$\sup_X |\phi_0^c - \phi_1^c| \leq \sup_X |\phi_1 - \phi_0|. \quad (38)$$

Once this is established the lemma follows from the dominated convergence theorem. To see that (38) holds, let  $y \in X$ . By compactness and continuity there is  $x_y \in X$  such that

$$\phi_0^c(y) = \sup_{x \in X} -c(x, y) + \phi_0(x) = -c(x_y, y) - \phi_0(x_y).$$

By construction

$$\phi_1^c(y) = \sup_{x \in X} -c(x, y) + \phi_1(x) \geq -c(x_y, y) - \phi_1(x_y).$$

We get

$$\phi_0^c(y) - \phi_1^c(y) \leq \phi_1(x_y) - \phi_0(x_y) \leq \sup_X |\phi_1 - \phi_0|.$$

By interchanging the roles of  $\phi_0$  and  $\phi_1$  we get

$$\phi_1^c(y) - \phi_0^c(y) \leq \sup_X |\phi_1 - \phi_0|$$

and hence that (38) holds.  $\square$

*Proof of Theorem 6.* We want to apply the Gärtner-Ellis theorem. As  $\chi = M_1(X)$  is compact, tightness of  $\Gamma^{(N)}$  holds automatically. By Lemma 5

$$\lim_{N \rightarrow \infty} \frac{1}{r_N} \log \Lambda_{\Gamma_k^{(N)}}(r_N \phi) = \xi(-\phi).$$

Further,  $\xi$  is finite valued since  $\phi^c$  is continuous, and hence bounded, for any  $\phi \in C(X)$ . By Lemma 6,  $\xi$  is continuous. Finally, by Corollary 2,  $\xi$  is Gateaux differentiable. As  $W^2(-\cdot, dx)$  is the Legendre transform of  $\xi$ , and hence  $W^2(\cdot, dx)$  is the Legendre transform of  $\xi(-\cdot)$ , the theorem follows from the Gärtner-Ellis theorem.  $\square$

### 3.2 A Thermodynamic Interpretation and Reduction to the Zero Temperature Case

The proof of Theorem 4 is based on a result on large deviation principles for Gibbs measures. Because of this we explain in this section how  $\{\mu_\beta^{(N)}\}$  can be seen as the Gibbs measures of certain thermodynamic systems. If we introduce the  $N$ -particle Hamiltonian

$$H^{(N)}(x_1, \dots, x_N) = -\frac{1}{k} \log \text{perm}(\Psi_{p_i}(x_j))$$

we may write  $\mu_\beta^{(N)}$  on the form

$$\mu_\beta^{(N)} = e^{-\beta H^{(N)}} d\mu_0^{\otimes N}.$$

This means  $\mu_\beta^{(N)}$  admits a thermodynamic interpretation as the *Gibbs measure*, or *canonical ensemble*, of the system determined by the Hamiltonian  $H^{(N)}$  and the background measure  $\mu_0$ . In this interpretation  $\mu_\beta^{(N)}$  is the equilibrium state of the system when the temperature is assumed fixed at  $Temp = 1/\beta$  and Theorem 6 is describing the zero-temperature limit. Theorem 4 will follow from Theorem 6 and a theorem on equi-continuous and uniformly bounded Hamiltonians. To state that theorem we need to define what it means for the family  $\{\frac{H^{(N)}}{N}\}$  to be equi-continuous. Let  $d(\cdot, \cdot)$  be the distance function induced by the standard Riemannian metric on  $X$ . This defines distance functions,  $d^{(N)}(\cdot, \cdot)$ , on  $X^N$  given by

$$d^{(N)}(x, y) = d^{(N)}(x_1, \dots, x_N, y_1, \dots, y_N) = \frac{1}{N} \inf_{\sigma} \sum_i d(x_i, y_{\sigma(i)}) \quad (39)$$

where the infimum is taken over all permutations  $\sigma$  of the set  $\{1, \dots, N\}$ . We will say that the family of functions  $\frac{H^{(N)}}{N}$  on  $X^N$  is (uniformly) equi-continuous if for every  $\epsilon > 0$  there is  $\delta > 0$  such that for all  $N$

$$\left| \frac{1}{N} H^{(N)}(x) - \frac{1}{N} H^{(N)}(y) \right| \leq \epsilon \quad (40)$$

whenever  $d^{(N)}(x, y) \leq d$ . Before we move on to state the Theorem 8 we prove the following well known lemma.

**Lemma 7.** *Let  $x = (x_1, \dots, x_N) \in X^N$  and  $y = (y_1, \dots, y_N) \in X^N$ . Then (39) is the optimal transport cost with respect to the cost function  $d(\cdot, \cdot)$ , of transporting the measure  $\delta^{(N)}(x) = \frac{1}{N} \sum \delta_{x_i}$  to the measure  $\delta^{(N)}(y) = \frac{1}{N} \sum \delta_{y_i}$ .*

*Proof.* We need to prove that

$$(39) = \inf_{\gamma} \int_{X \times X} d(x, y) \gamma \quad (41)$$

where the infimum is taken over all  $\gamma \in \mathcal{M}_1(X \times X)$  with first and second marginal given by  $\delta^{(N)}(x)$  and  $\delta^{(N)}(y)$  respectively. We will refer to any  $\gamma \in \mathcal{M}_1(X \times X)$  satisfying this as a feasible transport plan. The conditions on the marginals imply that any feasible transport plan is supported on the intersection of the sets  $\{x_i\} \times X$  and  $X \times \{y_i\}$ , in other words on the set  $\{x_i\} \times \{y_i\}$ . We conclude that the set of feasible transport plans is given by

$$\left\{ \sum_{i,j} a_{ij} \delta_{(x_i, y_j)} : a_{ij} \geq 0, \sum_i a_{ij} = 1/N, \sum_j a_{ij} = 1/N \right\}, \quad (42)$$

in other words a polytope in  $\mathcal{M}_1(X \times X)$ . It follows that the infimum in (41) is attained on one or more of the vertices of (42). Moreover, any permutation,  $\sigma$ , of  $N$  elements induce a feasible transport plan

$$\gamma_{\sigma} = \frac{1}{N} \sum_i \delta_{(x_i, y_{\sigma(i)})}$$

with transport cost

$$\int_{X \times X} d(x, y) \gamma_{\sigma} = \frac{1}{N} \sum_i d(x_i, y_{\sigma(i)}).$$

It is easy to verify that any vertex of (42) occur as  $\gamma_{\sigma}$  for some permutation  $\sigma$ . This proves the lemma.  $\square$

Note that this lemma implies that if we equip  $\mathcal{M}_1(X)$  with the Wasserstein 1-metric, which metricizes the weak\* topology on  $\mathcal{M}_1(X)$ , then the distance function defined in (39) makes the embeddings

$$\delta^{(N)} : X^N \hookrightarrow \mathcal{M}_1(X)$$

isometric embeddings.

**Theorem 8** ([16]). *Assume  $X$  is a compact manifold,  $\mu_0 \in \mathcal{M}_1(X)$ ,  $\{\frac{H^{(N)}}{N}\}$  is a uniformly bounded and equi-continuous family of functions on  $X^N$  and  $\beta_N$  is a sequence of numbers tending to infinity. Assume also that*

$$\left( \delta^{(N)} \right)_* e^{-\beta_N H^{(N)}} d\mu_0^{\otimes N}$$

satisfies a Large Deviation Principle with rate  $N\beta_N$  and rate function  $E$ . Then, for any  $\beta \in \mathbb{R}$ ,

$$\left(\delta^{(N)}\right)_* e^{-\beta H^{(N)}} d\mu_0^{\otimes N}$$

satisfies a Large Deviation Principle with rate  $N$  and rate function  $\beta E + Ent_{\mu_0}$ .

For completeness, we will include a proof of Theorem 8 here. It will be based on the following

**Proposition 1** ([16]). *Assume  $X$  is a compact manifold,  $\mu_0 \in \mathcal{M}_1(X)$ ,  $\beta \in \mathbb{R}$ ,  $\{\frac{H^{(N)}}{N}\}$  is a family of functions on  $X^N$ . Assume also that there is a functional  $E$  on  $\mathcal{M}_1(X)$  satisfying*

$$\sup_{X^N} \left| \frac{H^{(N)}}{N} - E \circ \delta^{(N)} \right| \rightarrow 0 \quad (43)$$

as  $N \rightarrow \infty$ . Then

$$\left(\delta^{(N)}\right)_* e^{-\beta H^{(N)}} \mu_0^{\otimes N}$$

satisfies a Large Deviation Principle with rate  $N$  and rate function  $\beta E + Ent_{\mu_0}$ .

*Proof.* Let  $\mu \in \mathcal{M}_1(X)$  and  $B_d(\mu)$  be the ball of (Wasserstein-1) radius  $d$  centred at  $\mu$  and

$$B_d^{(N)}(\mu) = (\delta^{(N)})^{-1}(B_d(\mu)) \subset X^N.$$

Using (43) we get

$$\begin{aligned} & \lim_{d \rightarrow 0} \liminf_{N \rightarrow \infty} -\frac{1}{N} (\delta^{(N)})_* e^{-\beta H^{(N)}} \mu_0^{\otimes N}(B_d(\mu)) \\ &= \lim_{d \rightarrow 0} \liminf_{N \rightarrow \infty} -\frac{1}{N} \log \int_{B_d^{(N)}(\mu)} e^{-\beta H^{(N)}(x)} d\mu_0^{\otimes N} \\ &= \lim_{d \rightarrow 0} \liminf_{N \rightarrow \infty} -\frac{1}{N} \log \int_{B_d^{(N)}(\mu)} e^{-\beta N(E \circ \delta^{(N)}(x) + o(1))} d\mu_0^{\otimes N} \\ &= \beta E(\mu) + \lim_{d \rightarrow 0} \liminf_{N \rightarrow \infty} -\frac{1}{N} \log \int_{B_d^{(N)}(\mu)} d\mu_0^{\otimes N}. \end{aligned} \quad (44)$$

and similarly with  $\liminf$  replaced by  $\limsup$  (here  $o(1) \rightarrow 0$  uniformly in  $x$  as  $N \rightarrow \infty$ ). By Sanov's theorem  $(\delta^{(N)})_* \mu_0^{\otimes N}$  satisfies a large deviation principle with rate  $N$  and rate function  $Ent_{\mu_0}$ . Hence, by Lemma 4, the second term in (44) is  $Ent_{\mu}(\mu)$ . Using Lemma 4 again, this proves the proposition.  $\square$

It turns out that in the compact setting, under the assumptions of uniform boundedness and equi-continuity, the assumption of convergence in Proposition 1 always holds for some functional  $U$  on  $\mathcal{M}_1(X)$ .

**Lemma 8.** Assume  $X$  is a compact manifold,  $\mu_0 \in \mathcal{M}_1(X)$  and  $\{\frac{H^{(N)}}{N}\}$  is a uniformly bounded and equi-continuous family of functions on  $X^N$ . Then there is a function  $U$  on  $\mathcal{M}_1(X)$  such that, after possibly passing to a subsequence,

$$\sup_{X^N} \left| \frac{H^{(N)}(x)}{N} - U \circ \delta^{(N)}(x) \right| \rightarrow 0 \quad (45)$$

as  $N \rightarrow \infty$ .

*Proof.* Using the embeddings  $\delta^{(N)} : X^n \hookrightarrow \mathcal{M}_1(X)$  the functions  $H^{(N)}$  define a sequence of functionals,  $\mathcal{H}^{(N)}$ , defined on the subspaces  $\delta^{(N)}(X^N) \subset \mathcal{M}_1(X)$ . By a standard procedure (we will explain it below) it is possible to define an equi-continuous family of extensions,  $\{U^{(N)}\}$ , of  $\frac{\mathcal{H}^{(N)}}{N}$  on  $\mathcal{M}_1(X)$ . By Arzelà-Ascoli theorem  $U^{(N)}$ , after possibly passing to a subsequence, will converge to a functional  $U$  satisfying (45). We may define the extensions  $U^{(N)}$  in the following way: Note that by assumption the functions  $\frac{H^{(N)}}{N}$  all satisfy the same modulus of continuity,  $\omega$ . We define  $U^{(N)} : \mathcal{M}_1(X) \rightarrow \mathbb{R}$  as

$$U^{(N)}(\mu) = \inf_{\nu \in \delta^{(N)}(X^N)} \frac{\mathcal{H}^{(N)}(\nu)}{N} + \omega(d(\mu, \nu))$$

where  $d(\cdot, \cdot)$  is the Wasserstein 1-distance on  $\mathcal{M}_1(X)$ . It follows from the definition of moduli of continuity that  $U^{(N)} = \frac{\mathcal{H}^{(N)}}{N}$  on  $\delta^{(N)}(X^N)$ . As  $\mathcal{M}_1(X)$  is compact we may take  $\omega$  to be sub-additive. It follows that the function  $\omega(d(\mu, \cdot))$  satisfies  $\omega$  as modulus of continuity. This means  $U^{(N)}$ , being a supremum of functions satisfying  $\omega$ , also satisfy  $\omega$ . In particular the family  $\{U^{(N)}\}$  is equi-continuous.  $\square$

We can now prove Theorem 8.

*Proof of Theorem 8.* As above, let  $B_d^{(N)}(\mu) = (\delta^{(N)})^{-1}(B_d(\mu)) \subset X^N$ , where  $B_d(\mu)$  is the ball in  $\mathcal{M}_1(X)$  centered at  $\mu$  with radius  $d$ . By the assumed Large Deviation Principle and Lemma 4, for any  $\mu \in \mathcal{M}_1(X)$ ,

$$E(\mu) = \lim_{d \rightarrow 0} \liminf_{N \rightarrow \infty} -\frac{1}{N\beta_N} \log \int_{B_d^{(N)}(\mu)} e^{-\beta_N H^N} \mu_0^{\otimes N}.$$

On the other hand, by Lemma 8 there is a function  $U$  on  $\mathcal{M}_1(X)$  such that, after possibly passing to a subsequence, (45) holds. This means

$$\begin{aligned} E(\mu) &= \lim_{\delta \rightarrow 0} \liminf_{N \rightarrow \infty} -\frac{1}{N\beta_N} \log \int_{B_d^{(N)}(\mu)} e^{-N\beta_N(U \circ \delta^{(N)} + o(1))} d\mu_0^{\otimes N} \\ &= U(\mu) + \lim_{\delta \rightarrow 0} \liminf_{N \rightarrow \infty} -\frac{1}{N\beta_N} \log \int_{B_d^{(N)}(\mu)} d\mu_0^{\otimes N} \\ &= U(\mu). \end{aligned} \quad (46)$$

where the second term in (46) is zero by Sanov's theorem. This means  $E = U$  and the theorem now follows from Proposition 1.  $\square$



### 3.3 Proof of Theorem 4

To use Theorem 8 we need to verify that the family  $\{H^{(N)}\}$  is equi-continuous. We will use the following two lemmas

**Lemma 9.** *The functions in  $P(X)$  are Lipschitz with the Lipschitz constant  $L = 1$ .*

*Proof.* As the diameter of  $X$  is 1 we get that the set

$$\{d(\cdot, y)^2/2 : y \in X\}$$

is Lipschitz with the Lipschitz constant  $L = 1$ . Now, assume  $\phi \in P(X)$  and  $x_1, x_2 \in X$ . By definition

$$\phi(x) = \sup_{y \in X} -d(x, y)^2/2 - \phi^c(y).$$

for all  $x$ . By compactness and continuity there is  $y_1$  such that

$$\phi(x_1) = -d(x_1, y_1)^2/2 - \phi^c(y_1).$$

We have

$$\begin{aligned} \phi(x_2) &\geq -d(x_2, y_1)^2/2 - \phi^c(y_1) = \phi(x_1) - (d(x_2, y_1)^2/2 - d(x_1, y_1)^2/2) \\ &\geq \phi(x_1) - d(x_1, x_2). \end{aligned}$$

By interchanging the roles of  $x_1$  and  $x_2$  we get

$$\phi(x_1) \geq \phi(x_2) - d(x_1, x_2)$$

and hence

$$|\phi(x_1) - \phi(x_2)| \leq d(x_1, x_2). \quad \square$$

We say that a function,  $\Phi$ , on  $\mathbb{R}^n$  is  $\lambda$ -convex if  $\Phi - \lambda \frac{x^2}{2}$  is convex.

**Lemma 10.** *Assume  $\Phi_\alpha$  is a family of functions on  $\mathbb{R}^n$  parametrized over some set  $A$ . Assume that for all  $\alpha \in A$ ,  $\Phi_\alpha$  is  $\lambda$ -convex. Let  $\sigma$  be a probability measure on  $A$ . Then*

$$\log \int e^{\Phi_\alpha} d\sigma(\alpha)$$

*is  $\lambda$ -convex.*

*Proof.* Assume first  $\lambda = 0$ . By the convexity of  $\Phi_\alpha$  in  $x$  and Hölder's inequality we get

$$\begin{aligned} \int_A e^{\Phi_\alpha(tx_1 + (1-t)x_0)} d\sigma(\alpha) &\leq \int_A e^{t\Phi_\alpha(x_1) + (1-t)\Phi_\alpha(x_0)} d\sigma(\alpha) \\ &\leq \left( \int_A e^{\Phi_\alpha(x_1)} d\sigma(\alpha) \right)^t \left( \int_A e^{\Phi_\alpha(x_0)} d\sigma(\alpha) \right)^{(1-t)} \end{aligned}$$

and hence, taking the logarithm of both sides of this inequality,

$$\begin{aligned} & \log \int_A e^{\Phi_\alpha(tx_1+(1-t)x_0)} d\sigma(\alpha) \\ & \leq t \log \int_A e^{\Phi_\alpha(x_1)} d\sigma(\alpha) + (1-t) \log \int_X e^{\Phi_\alpha(x_0)} d\sigma(\alpha). \end{aligned}$$

For the general case, note that

$$\log \int_A e^{\Phi_\alpha(x)} d\sigma(\alpha) - \lambda \frac{x^2}{2} = \log \int_A e^{\Phi_\alpha(x) - \lambda x^2/2} d\sigma(\alpha)$$

which is convex by the case considered above.  $\square$

We get

**Corollary 3.** *The normalized energy functions*

$$\{H^{(N)}/N : k \in \mathbb{N}\}$$

*is an equi-continuous family (in the sense of (40)).*

*Proof.* We claim that

$$c_p^{(N)} = \frac{1}{k} \log \sum_{m \in \mathbb{Z}^{n+p}} e^{-k|x-m|^2/2} \in P(X) \quad (47)$$

for all  $p \in X$  and  $k \in \mathbb{N}$ . To prove the claim it suffices to prove that (47) is  $-1$ -convex. This follows from Lemma 10 as  $-|x-m|^2/2$  is  $-1$ -convex for all  $m \in \mathbb{R}^n$ . Further, fixing all but one variable we get a function on  $X$  given by

$$\begin{aligned} x & \mapsto H^{(N)}(x_1, \dots, x_{i-1}, x, x_{i+1}, \dots, x_n) \\ & = \frac{1}{k} \log \sum_{\sigma} e^{-kc_{p_{\sigma(i)}}^{(N)}(x)} \prod_{j \neq i} e^{-kc_{p_{\sigma(j)}}^{(N)}(x_j)} \end{aligned}$$

By Lemma 10 this function is in  $P(X)$ . By Lemma 9 it satisfies the Lipschitz constant 1. This means, if  $x = (x_1, \dots, x_N)$  and  $y = (y_1, \dots, y_N)$  are points in  $X^N$ , that

$$\begin{aligned} & \left| \frac{1}{N} H^{(N)}(x_1, \dots, x_N) - \frac{1}{N} H^{(N)}(y_1, \dots, y_N) \right| \\ & \leq \frac{1}{N} \sum_i \left| H^{(N)}(x_1, \dots, x_{i-1}, y_i, \dots, y_N) - H^{(N)}(x_1, \dots, x_i, y_{i+1}, \dots, y_N) \right| \\ & \leq \sum_i d(x_i, y_i). \end{aligned} \quad (48)$$

As  $H^{(N)}$  is symmetric we may reorder  $\{x_i\}$  so that

$$\sum_i d(x_i, y_i) = \inf_{\sigma} \sum_i d(x_i, y_{\sigma(i)})$$

and hence the right hand side of (48) equals  $d^{(N)}(x, y)$ . This implies  $H^{(N)}/N$  is equi-continuous in the sense of (40).  $\square$

*Proof of Theorem 4.* By Theorem 6 and Theorem 8 we only need to verify that the family  $\{H^{(N)}/N\}$  is uniformly bounded and equi-continuous. The latter was proved in Corollary 3. To see that  $\{H^{(N)}/N\}$  is uniformly bounded recall that in the proof of Theorem 6 we proved that  $-\frac{1}{k} \log \Psi_p^{(N)}(x) \rightarrow d(x, p)/2$  uniformly in  $x$  and  $p$ . Since  $d(\cdot, \cdot)$  is bounded on  $X \times X$  we get that there is constants  $c, C \in \mathbb{R}$  such that, for all but finitely many  $N$ ,

$$c \leq \frac{1}{k} \log \Psi_p^{(N)}(x) \leq C \quad (49)$$

for all  $x, p$ . As the functions  $\{\frac{1}{k} \log \Psi_p^{(N)}\}$  are bounded on  $X$  and there is only finitely many functions for each  $N$ , we may choose  $c$  and  $C$  such that (49) holds for all  $N$ . We get

$$H^{(N)}(x)/N = \frac{1}{kN} \log \sum_{\sigma} \prod_i e^{\log \Psi_{p_i}(x)} \leq \frac{1}{kN} \log \sum_{\sigma} \prod_i e^{kC} = \frac{\log N!}{kN} + C$$

and

$$H^{(N)}(x)/N = \frac{1}{kN} \log \sum_{\sigma} \prod_i e^{\log \Psi_{p_i}(x)} \geq \frac{1}{kN} \log \prod_i e^{kc} = c$$

for all  $N$  and  $x \in X^N$ . This proves the theorem.  $\square$

## 4 The Rate Function and its relation to Monge-Ampère equations

In this section we will show how the rate function,  $G$ , in Theorem 4 is related to Monge-Ampère equations. More precisely, we will establish a variational approach to equation (1) and then show that, under a certain condition, the minimizers of the  $G$  are the Monge-Ampère measures of solutions to (1) (see Lemma 13). This will allow us to finish the proof of Theorem 1.

### 4.1 The Variational Approach to Equation (1)

In the variational approach to equation (1) it is convenient to consider its normalized version:

$$\text{MA}(\phi) = \frac{e^{\beta\phi} \mu_0}{\int_X e^{\beta\phi} d\mu_0}. \quad (50)$$

We see that this equation is invariant under the action of  $\mathbb{R}$  on  $P(X)$  given by

$$C \mapsto (\phi \mapsto \phi + C). \quad (51)$$

Now, we will say that an equation admits a unique solution modulo  $\mathbb{R}$  if, for any two solutions  $\phi_1, \phi_2 \in C(X)$ ,  $\phi_1 - \phi_2$  is constant. It is easy to verify that (1) admits a unique solution if and only if (50) admits a unique solution modulo  $\mathbb{R}$ . We will consider a certain energy functional (the analog of the Ding functional

in complex geometry) whose stationary points correspond to weak solutions of (1). For given data  $(\mu_0, \beta)$  this energy functional has the form

$$F(\phi) = \xi(\phi) + \frac{1}{\beta} I_{\mu_0}(\beta\phi).$$

where  $I_{\mu_0}$  is defined as

$$I_{\mu_0}(\phi) = \log \int_X e^\phi d\mu_0.$$

**Lemma 11.** *Let  $\beta \neq 0$ . The functional  $I_{\mu_0}$  is Gateaux differentiable and*

$$dI_{\mu_0}|_\phi = \frac{e^\phi \mu_0}{\int_X e^\phi d\mu_0}.$$

*Consequently,  $F$  is Gateaux differentiable and  $\phi$  is a stationary point of  $F$  if and only if  $\phi$  is a weak solution (in the sense of Section 2.2) to (1).*

*Proof.* Let  $v \in C(X)$ . As  $v$  is bounded an application of the dominated convergence theorem gives

$$\begin{aligned} \frac{d}{dt}|_{t=0} I(\phi + tv) &= \frac{\frac{d}{dt}|_{t=0} \int_X e^{\phi+tv} d\mu_0}{\int_X e^\phi d\mu_0} \\ &= \frac{\int_X \frac{d}{dt}|_{t=0} e^{\phi+tv} d\mu_0}{\int_X e^\phi d\mu_0} \\ &= \frac{\int_X v e^\phi d\mu_0}{\int_X e^\phi d\mu_0}, \end{aligned}$$

proving the first two statements of the lemma. By Corollary 2,  $\xi$  is differentiable and  $d\xi|_\phi = -\text{MA}(\phi)$ . This means  $F$  is Gateaux differentiable and

$$dF|_\phi = -\text{MA}(\phi) + \frac{e^\phi \mu_0}{\int_X e^\phi d\mu_0}$$

proving the last statements of the lemma. □

## 4.2 The Minimizers of the Gibbs Free Energy

We will use the following well know property of the relative entropy function in the proof of Lemma 13.

**Lemma 12.** *Let  $\mu \in \mathcal{M}_1(X)$  and  $\phi \in C(X)$ . Then*

$$I_{\mu_0}(\phi) + \text{Ent}_{\mu_0}(\mu) \geq \int_X \phi d\mu \tag{52}$$

*with equality if and only if  $\mu = dI_{\mu_0}|_\phi$ .*

*Proof.* Assume first that  $\mu$  is absolutely continuous with respect to  $\mu_0$  and  $\mu_0$  is absolutely continuous with respect to  $\mu$ . By Jensen's inequality

$$\begin{aligned} I_{\mu_0}(\phi) &= \log \int_X e^{\phi} \frac{\mu_0}{\mu} d\mu \\ &\geq \int_X \phi d\mu - \int_X \log \frac{\mu}{\mu_0} d\mu \\ &= \int_X \phi d\mu - Ent_{\mu_0}(\mu) \end{aligned}$$

with equality if and only if  $e^{\phi} \frac{\mu_0}{\mu}$  is constant, or, equivalently,  $\mu$  is proportional to  $e^{\phi} \mu_0$ . As  $\mu$  is a probability measure this means

$$\mu = \frac{e^{\phi} \mu_0}{\int_X e^{\phi} d\mu_0} = dI|_{\phi}$$

proving the lemma in this special case. If  $\mu$  is not absolutely continuous with respect to  $\mu_0$  then  $Ent_{\mu_0}(\mu) = +\infty$  and the equality holds trivially. Finally, when  $\mu$  is absolutely continuous with respect to  $\mu_0$  but  $\mu_0$  is not absolutely continuous with respect to  $\mu$ , then replacing  $\mu_0$  by  $\chi \mu_0$ , where  $\chi$  is the characteristic function of the support of  $\mu$  doesn't change the right hand side of (52). Since

$$I_{\mu_0}(\phi) \geq \log \int e^{\phi} \chi d\mu_0$$

this reduces this case to the case when  $\mu_0$  is absolutely continuous with respect to  $\mu$ .  $\square$

We can now prove Lemma 13.

**Lemma 13.** *Assume  $\beta \neq 0$ ,  $F$  admits a unique minimizer modulo  $\mathbb{R}$  and  $\phi_*$  is a minimizer of  $F$ . Then*

$$\mu_* = \text{MA}(\phi_*) \tag{53}$$

*is the unique minimizer of the rate function*

$$G(\mu) = \beta W^2(\mu, dx) + Ent_{\mu_0}(\mu) + C_{\mu_0, \beta}$$

*defined in Theorem 4.*

**Remark 3.** Note that  $\phi_1 - \phi_2 = C$  implies  $\phi_1^c - \phi_2^c = -C$  and hence

$$\text{MA}(\phi_1) = (\nabla^c \phi_1^c)_* dx = (\nabla^c \phi_2^c)_* dx = \text{MA}(\phi_2).$$

This means that, under the assumptions of Lemma 13,  $\mu_*$  is uniquely determined by (53).

*Proof of Theorem 13.* Note that by Corollary 2 and Lemma 12 we have, for all  $\mu \in M_1(X)$  and  $\phi \in C(X)$ , the two inequalities

$$W^2(\mu, dx) + \xi(\phi) \geq - \int \phi d\mu \quad (54)$$

$$Ent_{\mu_0}(\mu) + I_{\mu_0}(\phi) \geq \int \phi d\mu \quad (55)$$

where equality in (54) is characterized by

$$d\xi|_{\phi} = -MA(\phi) = -\mu \quad (56)$$

and equality in (55) is characterized by  $dI|_{\phi} = \mu$ . We will start with the case  $\beta > 0$ . Let  $\mu \in \mathcal{M}_1(X)$  and  $\phi_*$  be the minimizer of  $F$ . Applying (54) to the pair  $\mu$  and  $\phi_*$  and (55) to the pair  $\mu$  and  $\beta\phi_*$  we get

$$\begin{aligned} G(\mu) &= \beta W^2(\mu, dx) + Ent(\mu) \\ &\geq -\beta \int \phi_* d\mu - \beta \xi(\phi_*) + \int \beta \phi_* d\mu - I(\beta \phi_*) \\ &= -\beta \left( \xi(\phi_*) + \frac{1}{\beta} I(\beta \phi_*) \right) = -\beta F(\phi_*) \end{aligned}$$

with equality if and only if  $d\xi|_{\phi_*} = -MA(\phi_*) = -\mu$  and  $\mu = dI|_{\phi_*}$  which, since  $d\xi|_{\phi_*} + dI|_{\phi_*} = 0$ , is true if and only if  $\mu = MA(\phi_*)$ . For the case  $\beta < 0$ , let  $\mu \in \mathcal{M}_1(X)$ . By Corollary 2 we may take  $\phi$  to satisfy equality in (54) and hence (56). A similar application of (54) and (55) as above, keeping in mind that we have equality in (54), give

$$\begin{aligned} G(\mu) &= \beta W^2(\mu, dx) + Ent(\mu) \\ &\geq -\beta \int \phi d\mu - \beta \xi(\phi) + \int \beta \phi d\mu - I(\beta \phi) \\ &= -\beta \left( \xi(\phi) + \frac{1}{\beta} I(\beta \phi) \right) = -\beta F(\phi) \geq -\beta F(\phi_*). \end{aligned} \quad (57)$$

Moreover, equality in (58) holds if and only if  $\phi = \phi_*$ . But that means  $dI|_{\phi} = -d\xi|_{\phi} = \mu$ , hence we have equality in (57) as well. This implies  $G(\mu) \geq -\beta F(\phi_*)$  with equality if and only if  $\mu = MA(\phi_*)$ .  $\square$

### 4.3 Proof of Theorem 1 and Corollary 1

*Proof of Theorem 1.* Let  $\phi_*$  be the unique solution to (1). It follows that (50) admits a unique solution modulo  $\mathbb{R}$  and that  $\phi_*$  is a solution to (50). Now, we will use two results from the next chapter. Namely that any stationary point of  $F$  is a *smooth* solution to (50) (see Section 5.2) and that  $F$  always admit a minimizer (see Section 5.1). Under our assumptions, this implies  $F$  admits a unique minimizer modulo  $\mathbb{R}$  and that  $\phi_*$  is a minimizer of  $F$ . Using Lemma 13 we get that  $G$  admits the unique minimizer  $\mu_*$  satisfying  $\mu_* = MA(\phi_*)$ .

We want to prove that  $\Gamma^{(N)} \rightarrow \delta_{\mu_*}$  in the weak\* topology on  $\mathcal{M}_1(\mathcal{M}_1(X))$ . By the Portmanteau Theorem it suffices to verify that

$$\limsup_{N \rightarrow \infty} \Gamma^{(N)}(F) \leq \delta_{\mu_*}(F) \quad (59)$$

for all closed  $F \subset \mathcal{M}_1(X)$ . If  $\mu_* \in F$  then (59) holds trivially. Assume  $\mu_* \notin F$ . Recall that  $\mathcal{M}_1(X)$  is compact. This means the closed subset  $F$  is compact. Since  $G$  is lower semi-continuous there is  $\mu_F \in F$  such that  $\inf_F G = G(\mu_F)$ . As  $\mu_* \notin F$  is the unique point where  $G = \inf G = 0$  we get that  $G(\mu_F) = \inf_F G > 0$ . By the large deviation principle in Theorem 4

$$\limsup_{N \rightarrow \infty} \frac{1}{r_N} \log \Gamma^{(N)}(F) \leq -\inf_F G < 0.$$

As  $r_N \rightarrow \infty$  we get that  $\limsup \log \Gamma^{(N)}(F) = -\infty$  and  $\limsup \Gamma^{(N)}(F) = 0$ . This proves the theorem.  $\square$

*Proof of Corollary 1.* Equation (7) implies the first marginals of  $\mu_\beta^{(N)}$ ,

$$\int_{X^{N-1}} \mu_\beta^{(N)},$$

converges to  $\mu_*$  in the weak\* topology of  $M_1(X)$  (see Proposition 2.2 in [25]). Now,  $e^{\beta\phi_N}$  is the density with respect to  $\mu_0$  of the first marginal of  $\mu_\beta^{(N)}$ .

We claim that the collection  $\{\phi^{(N)} : k \in \mathbb{N}\}$  is equi-continuous and uniformly bounded. To see this, note that by Lemma 10,  $\phi^{(N)}$  is  $-1$ -convex and hence in  $P(X)$ . By Lemma 9 the functions  $\{\phi^{(N)}, k \in \mathbb{N}\}$  satisfy the Lipschitz constant  $L = 1$ . As

$$\int_X e^{\beta\phi_N} \mu_0 = \int_{X^N} \mu_\beta^{(N)} = 1$$

for all  $N$ , this means there are constants  $c, C \in \mathbb{R}$ , independent of  $N$ , such that  $c \leq \phi_N \leq C$ . This proves the claim. By the Arzelà-Ascoli theorem there is some function  $\phi_\infty \in C(X)$  such that

$$\phi_N \rightarrow \phi_\infty$$

uniformly. As

$$e^{\beta\phi_N} \mu_0 = \int_{X^{N-1}} \mu_\beta^{(N)} \rightarrow \mu_* = e^{\beta\phi_*} \mu_0$$

in the weak\* topology of  $M_1(X)$  we get that  $\phi_\infty = \phi_*$  almost everywhere with respect to  $\mu_0$ . As  $\mu_0$  has full support and  $\phi_\infty, \phi \in C(X)$ , this means  $\phi_\infty = \phi_*$ .  $\square$

## 5 Existence and Uniqueness of Solutions

In this section we will treat questions of existence and uniqueness of solutions to (1) for different data  $(\mu_0, \beta)$ . First of all we will prove that, for any data  $(\mu_0, \beta \neq 0)$ , (1) admit a weak solution. We will then explain how to reduce the problem of regularity to the case considered in [3], where the authors use Caffarelli's interior regularity theory for Monge-Ampère equations. In the last part of the section we treat uniqueness. We first prove the claim made in Remark 1, namely that as long as  $\beta > 0$  equation (1) admits at most one solution. Finally we prove Theorem 2 regarding  $\beta \in [-1, 0)$  and  $\mu_0 = \gamma$ .

### 5.1 Existence of Weak Solutions

First of all, Lemma 9 implies  $P(X)$  satisfies the following (relative) compactness property:

**Lemma 14.** *Let  $\{\phi_k\}$  be a sequence of functions in  $P(X)$  such that  $\inf_X \phi_k = 0$  for all  $k$ , then there is  $\phi \in C(X)$  such that, after passing to a subsequence,  $\phi_k \rightarrow \phi$  uniformly.*

*Proof.* By lemma 9,  $\{\phi_k\}$  are Lipschitz with a uniform Lipschitz constant. As  $X$  has finite diameter and  $\inf_X \phi_k = 0$  for all  $k$  this means  $\{\phi_k\}$  is also uniformly bounded, hence the lemma follows from the Arzelà-Ascoli theorem.  $\square$

**Lemma 15.** *Let  $\phi \in C(X)$  and*

$$F(\phi) = \xi(\phi) + \frac{1}{\beta} I_{\mu_0}(\beta\phi).$$

*Then*

$$F((\phi^c)^c) \leq F(\phi). \quad (60)$$

*Moreover, if  $\mu_0$  has full support, then equality holds in (60) if and only if  $\phi \in P(X)$ .*

*Proof.* Recall that  $\phi^c \in P(X)$ , and hence  $((\phi^c)^c)^c = \phi^c$  for all  $\phi \in C(X)$ . Also,  $(\phi^c)^c \leq \phi$  for all  $\phi \in C(X)$ . This means  $\xi(\phi) = \xi((\phi^c)^c)$  and

$$I_{\mu_0}((\phi^c)^c) = \frac{1}{\beta} \log \int_X e^{\beta(\phi^c)^c} d\mu_0 \leq \frac{1}{\beta} \log \int_X e^{\beta\phi} d\mu_0 = I_{\mu_0}(\phi). \quad (61)$$

and hence

$$F((\phi^c)^c) \leq F(\phi). \quad (62)$$

Assume  $\mu_0$  has full support. Then, if  $\phi \notin P(X)$  and hence  $(\phi^c)^c(x) < \phi(x)$  for some  $x \in X$ , then, as both  $(\phi^c)^c$  and  $\phi$  are continuous and  $\mu_0$  has full support, strict inequality holds in (61) and (62). This proves the lemma.  $\square$

**Lemma 16.** *Let  $\beta \in \mathbb{R} \setminus \{0\}$ . Then  $F$  admits a minimizer. In other words, (1) admits a weak solution.*



*Proof.* Recall that

$$F(\phi) = \xi(\phi) + \frac{1}{\beta} I(\beta\phi).$$

By the Dominated Convergence Theorem  $\frac{1}{\beta} I(\beta\phi)$  is continuous in  $\phi$ . By Lemma 6,  $\xi$  is continuous. This means  $F$  is continuous. Let  $\phi_k$  be a sequence such that  $F(\phi_k) \rightarrow \inf F$ . By Lemma 15 we may assume  $\phi_k \in P(X)$  for all  $k$ . As  $F$  is invariant under the action of  $\mathbb{R}$  given in (51) we may assume  $\phi_k$  satisfies  $\inf \phi_k = 0$  for all  $k$ . By Lemma 14, after possibly passing to a subsequence,  $\phi_k \rightarrow \phi$  for some  $\phi \in C(X)$ . By continuity  $F(\phi) = \lim_{k \rightarrow \infty} F(\phi_k) = \inf F$ , hence  $\phi$  is a minimizer of  $F$ .  $\square$

## 5.2 Regularity

In a number of papers (see [6], [7], [8]) Caffarelli developed a regularity theory for various types of weak solutions to Monge-Ampère equations. In particular, Caffarelli's theory applies to so called *Alexandrov solutions*. Recall that if  $f$  is a smooth function on  $\mathbb{R}^n$ , then a convex function  $\Phi$  on  $\mathbb{R}^n$  is an Alexandrov solution to the equation

$$\det(\Phi_{ij}) = f$$

if, for any borel measurable  $E \subset \Omega$ ,

$$\int_E f dx = \int_{\partial\Phi(E)} dx$$

where  $\partial\Phi(E)$  is the image of  $E$  under the multivalued gradient mapping, in other words

$$\partial\Phi(E) = \{y \in \mathbb{R}^n : \Phi(x) + \Phi^*(y) = \langle x, y \rangle \text{ for some } x \in E\}.$$

We have the following lemma:

**Lemma 17.** *Assume  $\mu_0$  is absolutely continuous with density  $f$  with respect to  $dx$ ,  $\beta \in \mathbb{R}$  and*

$$\text{MA}(\phi) = e^{\beta\phi} \mu_0. \tag{63}$$

*in the sense of Definition 1. Then  $\Phi = \phi \circ \pi + x^2/2$  is an Alexandrov solution to the equation*

$$\det(\Phi_{ij}) = e^{\beta(\Phi - x^2/2)} f \circ \pi \tag{64}$$

*on  $\mathbb{R}^n$ . Moreover,  $\Phi$  is proper.*

*Proof.* Assume  $E$  is a Borel measurable subset of  $\mathbb{R}^n$ . To prove the first point in the lemma we need to prove

$$\int_E e^{\beta(\Phi - x^2/2)} f \circ \pi dx = \int_{\partial\Phi(E)} dx.$$

Let  $C_0 = [0, 1]^n \subset \mathbb{R}^n$  and  $\{C_i\}$  be a collection of disjoint translates of  $C_0$  such that  $E \subset \cup C_i$ . Let  $E_i = E \cap C_i$ . We have

$$\int_E e^{\beta(\Phi - x^2/2)} f \circ \pi dx = \sum_i \int_{E_i} e^{\beta(\Phi - x^2/2)} f \circ \pi dx = \sum_i \int_{\pi(E_i)} e^{\beta\phi} f dx$$

and by (63)

$$\sum_i \int_{\pi(E_i)} e^{\beta\phi} f dx = \sum_i \int_{(\nabla^c \phi^c)^{-1}(\pi(E_i))} dx.$$

Now, we claim that  $\pi$  maps  $(\nabla\Phi^*)^{-1}(E_i)$  bijectively onto

$$(\nabla^c \phi^c)^{-1}(\pi(E_i))$$

for all  $i$ . To see this note that if  $y \in (\nabla\Phi^*)^{-1}(E_i)$ , then

$$\nabla^c \phi^c \circ \pi(y) = \pi \circ \nabla\Phi^*(y) \in \pi(E_i),$$

hence  $\pi(y) \in (\nabla^c \phi^c)^{-1}(\pi(E_i))$ . On the other hand, if  $y \in (\nabla^c \phi^c)^{-1}(\pi(E_i))$ , let  $\tilde{x}$  be the unique lift of  $\nabla^c \phi^c(y)$  in  $E_i$ . Moreover, let  $\tilde{y}$  be a lift of  $y$  in  $\mathbb{R}^n$ . Since  $\nabla^c \phi^c(y) = x$  we have  $\nabla\Phi^*(\tilde{y}) = \tilde{x} + m_0$  for some  $m_0 \in \mathbb{Z}^n$ . We have that

$$\pi^{-1}(y) = \{\tilde{y} + m : m \in \mathbb{Z}^n\}$$

and by (11)

$$\nabla\Phi^*(\tilde{y} + m) = \nabla\Phi^*(\tilde{y}) + m = \tilde{x} + m_0 + m.$$

We conclude that  $\nabla\Phi^*(\tilde{y} + m) \in E_i$  if and only if  $m = -m_0$  and then  $\nabla\Phi^*(\tilde{y} + m) = \tilde{x}$ . This means  $\pi$  maps  $(\nabla\Phi^*)^{-1}(E_i)$  bijectively onto  $(\nabla^c \phi^c)^{-1}(\pi(E_i))$  as claimed. We get

$$\sum_i \int_{(\nabla^c \phi^c)^{-1}(\pi(E_i))} dx = \sum_i \int_{(\nabla\Phi^*)^{-1}(E_i)} dx = \int_{(\nabla\Phi^*)^{-1}(E)} dx$$

where the second inequality holds since the sets  $(\nabla\Phi^*)^{-1}(E_i)$  are disjoint. Now, let  $\text{dom } \nabla\Phi^*$  be the set where  $\nabla\Phi^*$  is defined. We have

$$\begin{aligned} \text{dom } \nabla\Phi^* \cap \partial\Phi(E) &= \{y \in \mathbb{R}^n : \nabla\Phi^*(y) = x \text{ for some } x \in E\} \\ &= (\nabla\Phi^*)^{-1}(E). \end{aligned}$$

Since  $\Omega \setminus \text{dom } \nabla\Phi^*$  is a zero-set with respect to  $dx$  we have

$$\int_{(\nabla\Phi^*)^{-1}(E)} dx = \int_{\partial\Phi(E)} dx$$

which proves the first part of the lemma.

To see that  $\Phi$  is proper, note that since  $\phi$  is continuous it is bounded on  $X$ . Let  $C = \inf_X \phi$ . We get

$$\Phi(x) = \phi(\pi x) + \frac{x^2}{2} \geq C - 1 + |x|. \quad \square$$

**Lemma 18.** *Assume  $\mu_0$  is absolutely continuous with smooth density with respect to  $dx$  and  $\phi \in P(X)$  satisfies (1) in the sense of Definition 1. Then  $\phi$  is smooth.*

*Proof.* We refer to [3] (more precisely, step three in the proof of Theorem 1.1) where the authors explain why, by Caffarelli's regularity theory, proper Alexandrov solutions on  $\mathbb{R}^n$  to the equation

$$\det(\Phi_{ij}) = F(\Phi, x), \quad (65)$$

where  $F$  is smooth, are smooth. Strictly speaking the authors use an additional assumption of "finite energy", but the only way this is used is to guarantee properness of  $\Phi$ . By Lemma 17,  $\Phi = \phi \circ \pi + x^2/2$  is proper and satisfies (64) in the Alexandrov sense. As (64) is indeed a special case of (65) this proves the lemma.  $\square$

### 5.3 Uniqueness

We first prove the claim made in Remark 1.

**Theorem 9.** *Let  $\mu_0 \in \mathcal{M}_1(X)$  be absolutely continuous with smooth density with respect to  $dx$  and  $\beta > 0$ . Then (1) admits a unique solution.*

*Proof.* By Lemma 16 and Lemma 18 there always exist a solution to (1). To prove uniqueness it suffices to prove that the normalized equation (50) admits a unique solution modulo  $\mathbb{R}$ , in other words that  $F$  admits a unique minimizer modulo  $\mathbb{R}$ . Assume then  $\phi_0$  and  $\phi_1$  satisfies

$$F(\phi_0) = F(\phi_1) = \inf_{C(X)} F. \quad (66)$$

Let  $\phi_t = t\phi_1 + (1-t)\phi_0$ . Applying Lemma 10 with  $A = X$  and  $\Phi_\alpha(x) = \phi_x(\alpha)$  gives that

$$I_{\mu_0}(\phi_t) = \log \int_X e^{\phi_t} d\mu_0$$

is convex in  $t$ . Now,  $\xi(\phi_t)$  is convex in  $t$  by Lemma 3. This means  $F(\phi_t)$  is convex and hence, by (66), constant in  $t$ . It follows that  $I_{\mu_0}(\phi_t)$  is affine in  $t$ . However, if we let  $v = \frac{d}{dt}\phi_t = \phi_1 - \phi_0$ , then

$$\begin{aligned} \frac{d^2}{dt^2} I_{\mu_0}(\phi_t) &= \frac{d}{dt} \left( \frac{\int_X v e^{\phi_t} d\mu_0}{\int_X e^{\phi_t} d\mu_0} \right) \\ &= \frac{\int_X v^2 e^{\phi_t} d\mu_0 \int_X e^{\phi_t} d\mu_0 - \left( \int_X v e^{\phi_t} d\mu_0 \right)^2}{\left( \int_X e^{\phi_t} d\mu_0 \right)^2} \end{aligned} \quad (67)$$

Further, if we let  $\nu_t$  be the probability measure

$$\nu_t = \frac{e^{\phi_t} d\mu_0}{\int_X e^{\phi_t} d\mu_0}$$

and  $\hat{v}$  be the constant

$$\hat{v} = \int_X v \nu_t$$

then

$$(67) = \int_X v^2 \nu_t - \hat{v}^2 = \int_X v^2 \nu_t - 2\hat{v} \int_X v \nu_t + \hat{v}^2 = \int_X (v - \hat{v})^2 \nu_t.$$

In particular, since  $I_{\mu_0}(\phi_t)$  is affine in  $t$  we get that  $v = \hat{v}$ , hence that  $\phi_1 - \phi_0$  is constant. This proves the theorem.  $\square$

We now turn to the proof of Theorem 2. We will use

**Theorem 10** (The Prekopa Inequality [4], [15],[24]). *Let  $\phi : [0, 1] \times \mathbb{R}^n \rightarrow \mathbb{R}$  be a convex function. Define*

$$\hat{\phi}(t) = -\log \int_{\mathbb{R}^n} e^{-\phi(t,x)} dx.$$

Then, for all  $t \in \mathbb{R}$

$$\hat{\phi}(t) \leq t\hat{\phi}(1) + (1-t)\hat{\phi}(0)$$

with equality if and only if there is  $v \in \mathbb{R}^n$  and  $C \in \mathbb{R}$  such that

$$\phi(t, x) = \phi(0, x - tv) + tC.$$

*Proof of Theorem 2.* By Lemma 16 and Lemma 18 there always exist a solution to (3). Similarly as in the proof of Theorem 9, to prove uniqueness it suffices to prove that  $F$  admits a unique minimizer modulo  $\mathbb{R}$ . Assume  $\phi_0$  and  $\phi_1$  satisfies

$$F(\phi_0) = F(\phi_1) = \inf_{C(X)} F.$$

By Lemma 15 any minimizer of  $F$  is in  $P(X)$ , hence  $(\phi_0^c)^c = \phi_0$  and  $(\phi_1^c)^c = \phi_1$ . This means the following equation defines a curve in  $C(X)$  connecting  $\phi_0$  and  $\phi_1$ :

$$\phi_t = (t(\phi_1)^c + (1-t)(\phi_0)^c)^c. \quad (68)$$

Note that, as  $P(X)$  is convex and  $\phi_0^c, \phi_1^c \in P(X)$  we get  $t\phi_1^c + (1-t)\phi_0^c \in P(X)$  and

$$F(\phi_t) = \int_X t\phi_1^c + (1-t)\phi_0^c dx + \frac{1}{\beta} \log \int_X e^{\beta\phi_t} d\gamma. \quad (69)$$

The first term of this is affine in  $t$ . The second term is given by

$$\frac{1}{\beta} \log \int_X e^{\beta\phi_t} \sum_{m \in \mathbb{Z}^n} e^{|x-m|^2/2} dx = \frac{1}{\beta} \log \int_{\mathbb{R}^n} e^{\beta\phi_t \circ \pi - x^2/2} dx. \quad (70)$$

Let  $\Phi_t = \phi_t \circ \pi + x^2/2$ . By Lemma 1, since  $\phi_t$  is the  $c$ -transform  $t\phi_1^c + (1-t)\phi_0^c$ , we have

$$\Phi_t(x) = \sup_{y \in \mathbb{R}^n} \langle x, y \rangle - (t\phi_1^c + (1-t)\phi_0^c) \circ \pi(y) - \frac{y^2}{2}. \quad (71)$$

As

$$\langle x, y \rangle - (t\phi_1^c + (1-t)\phi_0^c) \circ \pi(y) - \frac{y^2}{2}$$

is affine in  $(t, x)$  we get that (71) is convex in  $(t, x)$ . It follows that, as long as  $\beta \in [-1, 0)$ , the exponent in (70),

$$\begin{aligned} \beta\phi_t \circ \pi(x) - x^2/2 &= \beta(\phi_t \circ \pi(x) + x^2/2) - (\beta+1)x^2/2 \\ &= \beta\Phi_t(x) - (\beta+1)x^2/2 \end{aligned}$$

is concave in  $(t, x)$ . We may then apply the Prekopa inequality to deduce that (70) and hence  $F(\phi_t)$  is convex in  $t$ . In particular, as  $\phi_0$  and  $\phi_1$  are minimizers of  $F$ , this means  $F(\phi_t) = F(\phi_0) = F(\phi_1)$  for all  $t \in [0, 1]$ . This implies (70) is affine in  $t$ . By the equality case in the Prekopa inequality

$$\beta\phi_1 \circ \pi(x) - x^2/2 = \beta\phi_0 \circ \pi(x-v) - (x-v)^2/2 + C$$

for some  $C \in \mathbb{R}$  and  $v \in \mathbb{R}^n$ . By noting that  $\phi_1 \circ \pi$  and  $\phi_0 \circ \pi(\cdot - v)$ , and hence

$$\beta\phi_1 \circ \pi - \beta\phi_0 \circ \pi(\cdot - v) = \langle \cdot, v \rangle + v^2/2 + C,$$

should descend to a function on  $X$  (in other words, they should be invariant under the action of  $\mathbb{Z}^n$ ), we get that  $v = 0$ . This means  $\phi_1 = \phi_0 + C$  which proves Theorem 2.  $\square$

## 6 Geometric Motivation

The original motivation for this project comes from the paper on statistical mechanics and birational geometry by Berman [2]. Berman introduces a thermodynamic approach to produce solutions to the complex Monge-Ampère equation

$$\text{MA}_{\mathbb{C}}(u) = e^{\beta u} \mu_0 \tag{72}$$

on a compact Kähler manifold  $M$ . The Monge-Ampère operator in (72) is defined as

$$(i\partial\bar{\partial}u + \omega_0)^n \tag{73}$$

where  $n$  is the complex dimension of  $M$  and  $\omega_0$  is a fixed Kähler-form on  $M$  representing the Chern class of a line bundle  $L$  over  $M$ . A solution,  $u$ , should be a real valued twice differentiable function on  $M$  satisfying  $i\partial\bar{\partial}u + \omega_0 > 0$ . As Berman's thermodynamic approach to this equation has served as an inspiration for us, we outline it here.

The metric,  $\omega_0$  determines, up to a constant, a metric on  $L$ . For each  $k > 0$ , let  $N = N_k = H^0(M, L)$ . By assumption on  $\omega_0$ ,  $L$  is ample and hence  $N_k \rightarrow \infty$  as  $k \rightarrow \infty$ . Let  $s_1, \dots, s_N$  be a basis of  $H^0(M, L)$ . Locally we may identify this basis with a collection of functions  $f_1, \dots, f_N$ . The map

$$(x_1, \dots, x_N) \mapsto \det(f_i(x_j))$$

determines a section,  $\det(s_1, \dots, s_N)$ , of the induced line bundle  $L^{\boxtimes N_k}$  over  $M_N$ . The metric on  $L$  induces a metric,  $\|\cdot\|$ , on this line bundle and

$$\|\det(s_1, \dots, s_N)\|^{2\beta/k} \mu_0 \quad (74)$$

determines a symmetric measure on  $M^N$ . Note that changing the basis of  $H^0(M, L)$  will give the same result up to a multiplicative constant. As long as this measure has finite volume we may normalize it to get a symmetric probability measure on  $M^N$ .

Now, Berman shows that if  $\beta > 0$  and the singularities of  $\mu_C$  are controlled in a certain way, then the point processes defined by (74) converge to the Monge-Ampère measure of a solution to (72). However, it should be stressed that when  $\beta < 0$  there is no guarantee that (74) has finite volume and can be normalized to a probability measure. This turns out to be a subtle property and in one of the most famous versions of equation (72), when  $M$  is a Fano manifold and  $\omega_0$ ,  $\mu$  and  $\beta$  are chosen so that solutions to (72) define Kähler-Einstein metrics of positive curvature, this reduces to a property of the manifold  $M$  which is conjectured to be equivalent to the existence of Kähler-Einstein metrics on  $M$  (see [17] for some progress on this). We will explain in Section 6.1 how equation (3) can be seen as the "push forward" to a real setting of a complex Monge-Ampère equation whose solution define Kähler-Einstein metrics of almost everywhere positive curvature. In that sense, the present project can be seen as an attempt to study one side of this complex geometric problem.

## 6.1 Equation (3) as the "Push Forward" of a Complex Monge-Ampère Equation

Let  $M = \mathbb{C}^n / (4\pi\mathbb{Z}^n + i\mathbb{Z}^n)$  and  $\theta$  be the function on  $\mathbb{C}^n$  defined as

$$\theta(z) = \sum_{m \in \mathbb{Z}^n} e^{-m^2/4 + izm/2}.$$

This is the classical  $\theta$ -function and it satisfies the following transformation properties:

$$\begin{aligned} \theta(z + 4\pi) &= \theta(z) \\ \theta(z + i) &= \theta(z) e^{iz/2 - 1/4}. \end{aligned}$$

In particular, the zero set of  $\theta$  defines the *theta divisor*,  $D$ , on  $M$  and, using certain trivializations of the line bundle associated to  $D$ ,  $\theta$  descends to a holomorphic section of this line bundle. This means  $\tau = i\partial\bar{\partial} \log |\theta|^2$  is a well-defined (1,1)-current on  $M$  and we may consider the twisted Kähler-Einstein equation

$$\text{Ric}(\omega) + \tau = \omega \quad (75)$$

on  $M$ , where  $\text{Ric}(\omega)$  denotes the Ricci curvature of  $\omega$ . The current  $\tau$  is supported on  $D$  so away from  $D$  this equation defines metrics of constant positive Ricci

curvature. Now, there is a standard procedure to rewrite (75) into a scalar equation of type (72). This process involves choosing a reference form  $\omega_0$  in the cohomology class of  $\tau$  and fixing a Ricci-potential of  $\omega_0$ ,  $F$ , such that

$$i\partial\bar{\partial}F = \text{Ric}(\omega_0) + \tau - \omega_0.$$

Choosing  $\omega_0 = \sum_i idz_i \wedge d\bar{z}_i$  and  $F = -y^2/2 + \log|\theta|^2$  gives the equation

$$\text{MA}_{\mathbb{C}}(u) = e^{-u-y^2/2}|\theta|^2\omega_0^n. \quad (76)$$

In other words, we arrive at equation (72) with the choices

$$\mu_{\mathbb{C}} = |\theta|^2 e^{-y^2/2} \omega_0^n$$

and  $\beta = -1$ . Now, let  $z = x + iy$  be the standard coordinates on  $M$  induced from  $\mathbb{C}^n$ . Let  $\rho : M \rightarrow X$  be the map  $z \mapsto y$ . If  $\phi$  is a twice differentiable function on  $X$  such that  $(\phi_{ij} + \delta_{ij})$  is strictly positive definite, then  $u(z) := \phi(y)$  defines a (rotationally invariant) twice differentiable function on  $M$  satisfying  $i\partial\bar{\partial}u + \omega_0 > 0$ . Moreover,

$$\rho_* \text{MA}_{\mathbb{C}}(u) = \text{MA}(\phi) \quad (77)$$

where  $\text{MA}(u)$  is the complex Monge-Ampère measure on  $M$  defined in (73) and  $\text{MA}(\phi)$  is the real Monge-Ampère measure on  $X$  defined in (2). Further, at the end of the next section we will prove

**Lemma 19.**

$$\rho_* \left( e^{-y^2/2} |\theta|^2 \omega_0^n \right) = \gamma \quad (78)$$

where  $dy$  is the uniform measure on  $X$ .

Since  $u$  is rotationally invariant we get that

$$\rho_* \left( \text{MA}_{\mathbb{C}}(u) - e^{-u-y^2/2} |\theta|^2 \omega_0^n \right) = \text{MA}(\phi) - e^{-\phi} \gamma$$

and this is the relation that makes us refer to equation (3) as the "push forward" of equation (76).

## 6.2 Permanental Point Processes as the Push Forward of Determinantal Point Processes

Here we will establish a connection between the permanental point processes defined in Section 1.2 and the determinantal point processes defined in Bermans framework. The connection is a consequence of a certain formula that relates integrals of determinants to permanents. This formula might be of independent interest and is given in the following lemma.

**Lemma 20.** Let  $(E, \mu)$  be a measure space. Let  $N \in \mathbb{N}$  and

$$\{F_{jk} : j = 1 \dots N, k = 1 \dots N\}$$

be a collection of complex valued functions on  $E$ , square integrable with respect to  $\mu$ , such that, for each  $j$

$$\int_E F_{jk} \overline{F_{jl}} d\mu = 0$$

if  $k \neq l$ . Then

$$\text{perm} \left( \int_E |F_{jk}|^2 d\mu \right) = \int_{E^N} |\det(F_{jk}(x_j))|^2 d\mu^{\otimes N}.$$

*Proof.* Now,

$$\begin{aligned} & \int_{E^N} |\det(F_{jk}(x_j))|^2 d\mu^{\otimes N} \\ &= \int_{E^N} \det(F_{jk}(x_j)) \overline{\det(F_{jk}(x_j))} d\mu^{\otimes N} \\ &= \int_{E^N} \left( \sum_{\sigma} (-1)^{\sigma} \prod_j F_{j\sigma(k)}(x_j) \right) \overline{\left( \sum_{\sigma'} (-1)^{\sigma'} \prod_j F_{j\sigma'(k)}(x_j) \right)} d\mu^{\otimes N} \\ &= \sum_{\sigma, \sigma'} (-1)^{\sigma + \sigma'} \prod_j \int_E F_{j\sigma(k)} \overline{F_{j\sigma'(k)}} d\mu \end{aligned} \quad (79)$$

By the orthogonality assumption on  $\{F_{jk}\}_k$ , the only contribution comes from terms where  $\sigma = \sigma'$ . We get

$$(79) = \sum_{\sigma} \prod_j \int_E |F_{j\sigma(k)}|^2 d\mu = \text{perm} \left( \int_E |F_{jk}|^2 d\mu \right). \quad \square$$

Before we examine its consequences for permanent point processes we illustrate two other applications. The first is given by a quick proof of the following well known formula related to Gram Determinants (see for example [14]):

**Corollary 4.** Let  $(E, \mu)$  be a measure space and

$$f_1, \dots, f_N \in L^2(\mu).$$

Then

$$\det \left( \int_E f_j \overline{f_k} d\mu \right) = \frac{1}{N!} \int_{E^N} |\det(f_k(x_j))|^2 d\mu^{\otimes N}. \quad (80)$$

*Proof.* Note that if  $A$  is an invertible  $N \times N$  matrix with determinant 1, then replacing  $\{f_1, \dots, f_N\}$  by  $\{\tilde{f}_1, \dots, \tilde{f}_N\}$  where  $\tilde{f}_i$  is defined by

$$(\tilde{f}_1, \dots, \tilde{f}_N) = (f_1, \dots, f_N)A$$



doesn't affect the formula (80). This means we may assume  $f_1, \dots, f_N$  satisfy

$$\int_E f_j \overline{f_k} d\mu = 0$$

if  $j \neq k$ . For each  $j, k \in \{1, \dots, N\}$ , let  $F_{jk} = f_k$ . We get that

$$\det \left( \int_E f_j \overline{f_k} d\mu \right) = \prod_k \int_E |f_k|^2 d\mu = \frac{1}{N!} \text{perm} \left( \int_E |F_{jk}|^2 d\mu \right)$$

and, applying Lemma 20, that

$$\begin{aligned} \det \left( \int_E f_j \overline{f_k} d\mu \right) &= \frac{1}{N!} \text{perm} \left( \int_E |F_{jk}|^2 d\mu \right) = \frac{1}{N!} \int_{E^N} |\det (F_{jk}(x_j))|^2 d\mu^{\otimes N} \\ &= \frac{1}{N!} \int_{E^N} |\det (f_k(x_j))|^2 d\mu^{\otimes N} \end{aligned}$$

proving the corollary.  $\square$

The second application of Lemma 20 is given by the following formula for the permanent of a matrix of non-negative real numbers.

**Corollary 5.** *Let  $(a_{jk})$  be an  $N \times N$ -matrix of non-negative real numbers. Then*

$$\text{perm}(a_{jk}) = \frac{1}{(2\pi)^N} \int_{[0, 2\pi]^N} |\det (\sqrt{a_{jk}} e^{ikx_j})|^2 dx_1 \cdots dx_N.$$

*Proof.* Let  $F_{jk} = \sqrt{a_{jk}} e^{ikx}$ . Then, for each  $j$ ,

$$\int_{[0, 2\pi]} F_{jk} \overline{F_{jl}} dx = \int_{[0, 2\pi]} a_{jk} e^{i(k-l)x} dx = \begin{cases} 2\pi a_{jk} & \text{if } l = k \\ 0 & \text{otherwise.} \end{cases}$$

Applying Lemma 20 gives

$$\begin{aligned} \text{perm}(a_{jk}) &= \frac{1}{(2\pi)^N} \text{perm} \int_{[0, 2\pi]} |F_{jk}|^2 dx \\ &= \frac{1}{(2\pi)^N} \int_{[0, 2\pi]^N} |\det (F_{jk}(x_j))|^2 dx_1 \cdots dx_N \\ &= \frac{1}{(2\pi)^N} \int_{[0, 2\pi]^N} |\det (\sqrt{a_{jk}} e^{ikx_j})|^2 dx_1 \cdots dx_N. \end{aligned}$$

which proves the corollary.  $\square$

To see how Lemma 20 connects permanental point processes to determinantal point processes, we will now look a bit closer on the point processes defined by Bermans framework when applied to the complex Monge-Ampère equation in Section 6.1. First of all,  $\omega_0 = \sum_i idz_i \wedge d\bar{z}_i$  represents the curvature class

of the theta divisor  $D$  on  $M$ . Elements in  $H^0(M, kD)$  may be represented by theta functions and a basis at level  $k \in \mathbb{N}$  is given by the set

$$\{\theta_p^{(k)} : p \in \frac{1}{k}\mathbb{Z}^n / \mathbb{Z}^n\} \quad (81)$$

where

$$\theta_p^{(k)} = \sum_{m \in \mathbb{Z}^n + p} e^{-km^2/4 + izkm/2}.$$

With respect to these trivializations the norm of  $\theta_p^{(k)}$  with respect to the metric on  $kD$  with curvature form  $k\omega_0$  may be written

$$\|\theta_p^{(k)}\|^2 = |\theta_p^{(k)}|^2 e^{-ky^2/2}.$$

Enumeration the points in  $\frac{1}{k}\mathbb{Z}^n / \mathbb{Z}^n$ ,  $\{p_1, \dots, p_N\}$  and using the standard coordinates  $(z_1, \dots, z_N) = (x_1 + iy_1, \dots, x_N + iy_N)$  on  $M^N$  allow us to write the determinant in (74) as

$$\left| \det \left( \theta_{p_i}^{(k)}(z_j) e^{-y_j^2/4} \right)_{j,l} \right|^2.$$

Now, recall that the real Monge-Ampère measure on  $X$  may be recovered as the push forward under the projection map,  $\rho : M \rightarrow X$ , of the complex Monge-Ampère measure on  $M$  (see equation (77)). Similarly, Lemma 20 will allow us to explicitly calculate the push forward of the measure

$$|\det(\theta_{p_i}^{(k)}(z_j) e^{-y_j^2/4})|^2 \omega_0^n$$

on  $M^N$  under the map  $\rho^{\times N} : M^N \rightarrow X^N$ . We get the following lemma, which is the key point of this section. It shows that the permanent point processes defined in Section 1.2 are the natural analog of the determinantal point processes defined by Bermans framework for complex Monge-Ampère equations.

**Lemma 21.** *Let  $dy$  be the uniform measure on  $X$ . Then*

$$(\rho^{\times N})_* |\det(\theta_{p_i}^{(k)}(z_j) e^{-y_j^2/4})|^2 \omega_0^n = \text{perm} \left( \Psi_{p_i}^{(N)}(y_j) \right) dy. \quad (82)$$

*Proof.* Let  $y = (y_1, \dots, y_N) \in X^N$ . The point  $y \in X^N$  defines a real torus,  $T_y$ , in  $M^N$

$$T_y = (\rho^{\times N})^{-1}(y) = \{x + iy : x \in (\mathbb{R}^n / 4\pi\mathbb{Z}^n)^N\}.$$

If we let  $dx$  be the measure on  $T_y$  induced by  $(\mathbb{R}^n)^N$ , then the density at  $y$  of the left hand side of (82) with respect to  $dy$  is given by the integral

$$\int_{T_y} |\det(\theta_{p_i}^{(k)}(z_j) e^{-y_j^2/4})|^2 dx. \quad (83)$$

For each  $j, l \in \{1, \dots, N\}$ , let  $F_{jl} : T_y \rightarrow \mathbb{C}$  be defined by

$$\begin{aligned} F_{jl}(x) &= \theta_{p_l}(x + iy_j)e^{-y_j^2/4} \\ &= \sum_{m \in \mathbb{Z}^n + p_l} e^{-km^2/4 + i(x + iy_j)km/2 - y_j^2/4} \\ &= \sum_{m \in \mathbb{Z}^n + p_l} e^{-k(m - y_j)^2/4 + ikmx/2} \end{aligned}$$

Now, when computing the integral

$$\begin{aligned} &\int_{T_y} F_{jl} \overline{F_{j'l'}} dx \\ &= \int_{T_y} \sum_{\substack{m \in \mathbb{Z}^n + p_l \\ m' \in \mathbb{Z}^n + p_{l'}}} e^{-k(m - y_j)^2/4 - k(m' - y_{j'})^2/4 + ik(m - m')x/2} dx \end{aligned} \quad (84)$$

the only contribution comes from the terms where  $m - m' = 0$ . If  $l \neq l'$ , then there are no such terms, in other words (84) = 0. If  $l = l'$  we are left with

$$(84) = (4\pi)^N \sum_{m \in \mathbb{Z} + p_l} e^{-k|y_j - m|^2/2} = (4\pi)^N \Psi_{p_l}^{(N)}(y_j).$$

Applying Lemma 20 gives

$$(83) = \int_{T_y} |\det(F_{jl}(x_j))|^2 dx = \text{perm} \left( \int |F_{jl}|^2 dx \right) = \text{perm}(\Psi_{p_l}(y_j))$$

proving the lemma.  $\square$

Finally, we show that Lemma 19 is a special case of this.

*Proof of Lemma 19.* Note that  $\theta = \theta_0^{(1)}$  and

$$\gamma = \sum_{m \in \mathbb{Z}^n} e^{-|y - m|^2/2} dy = \Psi_0^{(1)} dy.$$

This means (78) is the special case of (82) given by  $N = k = 1$ . Hence the lemma follows from Lemma 21.  $\square$

### 6.3 Approximations of Optimal Transport Maps

As mentioned in the introduction the point processes defined here can be used to produce explicit approximations of optimal transport maps. In optimal transport it is natural to consider a larger class of Monge-Ampère operators. Let  $\nu_0 \in \mathcal{M}_1(X)$  be absolutely continuous with respect to  $dx$ . Then  $\nu_0$  defines a Monge-Ampère operator  $\text{MA}_{\nu_0}$  on  $P(X)$  as

$$\text{MA}_{\nu_0}(\phi) = (\nabla^c \phi^c)_* \nu_0.$$

Solutions,  $\phi_*$ , to the inhomogenous Monge-Ampère equation

$$\text{MA}_{\nu_0}(\phi) = \mu_0 \quad (85)$$

determine optimal transport maps on  $X$  in the sense that  $T = \nabla^c \phi_*$  is the optimal *transport map* in the sense of Brenier (see [26]) from the source measure  $\mu_0$  to the target measure  $\nu_0$ .

The fact that the point processes defined in Section 1.2 are related to the standard  $\text{MA} = \text{MA}_{dx}$  is a consequence of the fact that

$$\frac{1}{N} \sum_{p \in \frac{1}{k} \mathbb{Z}^n / \mathbb{Z}^n} \delta_p \rightarrow dx$$

in the weak\*-topology. Redefining  $S^{(N)}$  in the following way will provide the generalisation we want: Let  $P^{(N)}$  be a collection of point sets with the property that  $|P^{(N)}| = N$  and

$$\frac{1}{N} \sum_{p \in P^{(N)}} \delta_p \rightarrow \nu_0.$$

As in the original definition, associate a wave function,  $\Psi_p^{(N)}$ , to each point  $p \in \cup P^{(N)}$

$$\Psi_{p_i}^{(K)} = \sum_{m \in \mathbb{Z}^n + p_i} e^{-|x-m|^2}$$

and, for each  $N$ , enumerate the points in  $P^{(N)}$

$$P^{(N)} = \{p_1, \dots, p_N\}.$$

We get

**Corollary 6.** *Let  $\mu_0, \nu_0 \in \mathcal{M}_1(X)$  be absolutely continuous and have smooth, strictly positive densities with respect to  $dx$  and  $\Psi_{p_i}^{(N)}$  be defined as above. Then*

$$\phi_N := \frac{1}{N} \log \int_{X^{N-1}} \text{perm} \left( \Psi_{p_i}^{(N)}(x_j) \right) d\mu^{\otimes N}$$

*converges uniformly to the unique, smooth, strictly convex solution of (85). Consequently, the associated gradient maps  $\nabla^c \phi_N$  converges uniformly to the unique optimal transport map transporting  $\mu_0$  to  $\nu_0$ .*

*Proof of Corollary 6.* First of all, the fact that the optimal transport map is smooth follow from Caffarelli's regularity theory for Monge-Ampère equations. We will not go through the argument as it is similar as in Section 5.2. Uniqueness is a basic result from optimal transport (see for example Theorem 2.4.7 in [26]). Now, to see that the convergence holds, consider the functionals,  $\{H^{(N)}\}$ , on  $\mathcal{M}_1(X)$  defined by

$$E^{(N)}(\mu) = \frac{1}{N} \int_{X^N} H^{(N)} d\mu^{\otimes N}.$$

Direct calculations give that they are continuous, convex, Gateaux differentiable and  $dE^{(N)}|_{\mu_0} = \Phi^{(N)}$ . We claim that

$$E^{(N)}(\mu) \rightarrow W^2(\mu, dx) \quad (86)$$

for all  $\mu \in \mathcal{M}_1(X)$ . To see this, note that by the proof of Theorem 8

$$\sup_{X^N} \left| \frac{1}{N} H^{(N)} - W^2(\cdot, dx) \circ \delta^{(N)} \right| \rightarrow 0$$

as  $N \rightarrow \infty$ . We get, since  $\{\frac{1}{N} H^{(N)}\}$  are uniformly bounded,

$$\begin{aligned} E^{(N)}(\mu) &= \int_{X^N} W^2(\cdot, dx) \circ \delta^{(N)} d\mu^{\otimes N} + o(1) \\ &= \int_{\mathcal{M}_1(X)} W^2(\cdot, dx) \left( \delta^{(N)} \right)_* \mu^{\otimes N} + o(1). \end{aligned} \quad (87)$$

where  $o(1) \rightarrow 0$  as  $N \rightarrow \infty$ . Now, it follows from Sanov's theorem that  $(\delta^{(N)})_* \mu^{\otimes N} \rightarrow \delta_\mu$  in the weak\*-topology on  $\mathcal{M}_1(\mathcal{M}_1(X))$ . Now, since  $X$  has finite diameter we get that the squared distance function on  $X$  can be bounded by a constant times the distance function. As the Wasserstein 1-metric metrizes the weak\* topology on  $\mathcal{M}_1(X)$  this implies that  $W^2(\cdot, dx)$  is continuous on  $\mathcal{M}_1(X)$ . We get that (87) converges to  $W^2(\mu, dx)$  as  $N \rightarrow \infty$ .

Further,  $W^2(\cdot, dx)$  is convex. By standard properties of convex functions  $dE^{(N)}|_{\mu_0}$  converges to a subgradient of  $W^2(\cdot, dx)$  at  $\mu_0$ . By standard properties of the Legendre Transform this means

$$\phi = \lim_{N \rightarrow \infty} \phi^{(N)} \quad (88)$$

satisfies  $d\xi|_\phi = MA(\phi) = \mu_0$ . This means  $\phi$  is smooth and  $\nabla^c \phi$  defines the optimal transport map transporting  $\mu_0$  to  $\nu_0$ . Now, let  $\Phi_N$  and  $\Phi$  be the images in  $P_{\mathbb{Z}^n}(\mathbb{R}^n)$  of  $\phi_N$  and  $\phi$  respectively. The convergence in (88) implies  $\Phi_N \rightarrow \Phi$  and, by standard properties of convex functions,  $\nabla \Phi^{(N)} \rightarrow \nabla \Phi$ . This means  $\nabla^c \phi_N \rightarrow \nabla^c \phi$  which proves the Corollary.  $\square$

## References

- [1] R.J. Berman *Statistical Mechanics of Permanents, Real Monge-Ampère Equations and Optimal Transport* arXiv:1302.4045v1 [Math.AP] 2013.
- [2] R.J. Berman *Kähler-Einstein metrics, Canonical Random Point Processes and Birational Geometry* arXiv:1307.3634v1 [Math.DG] 2013
- [3] R.J. Berman, B. Berndtsson *Real Monge-Ampère Equations and Kähler-Ricci Solitons on Toric Log Fano Varieties* arXiv:1207.6128v1 [Math.DG] 2012.

- [4] C. Borell. *Convex set functions in  $d$ -space* Period. Math. Hungar., Vol. 6. p 111-136. 1975.
- [5] Y. Brenier *Décomposition polaire et réarrangement monotone des champs de vecteurs*. C.R. Acad. Sci. Paris Sér. I Math., 305 (1987), 805-808.
- [6] L.A. Caffarelli. *Interior  $W^{2,p}$  estimates for solutions of the Monge-Ampère equation*. Ann. of Math. (2) 131 1990, no 1, 135-150.
- [7] L.A. Caffarelli. *A localization property of viscosity solutions to the Monge-Ampère equation and their strict convexity*. Ann. of Math. (2) 131 1990, no 1, 129-134.
- [8] L.A. Caffarelli. *Some regularity properties of solutions of Monge-Ampère equations* Comm. Pure Appl. Math. 44 no 8-9, 965-959. 1991.
- [9] L.A. Caffarelli. J.A. Viaclovsky. *On the Regularity of Solutions to Monge-Ampère Equations on Hessian Manifolds* Commun. in Partial Differential Equations, 26(11&12, 2339-2351 (2001).
- [10] S.-Y. Cheng S.-T. Yau *The Real Monge-Ampère Equation and Affine Flat Structures* Proceedings of the 1980 Beijing Symposium on Differential Geometry and Differential Equations, Vol. 1. 339-370, Science Press, Beijing, 1982.
- [11] D. Cordero-Erasquin *Sur le transport de mesures périodiques* C.R. Acad. Sci. Paris Sr. I Math. 329 (1999), 199–202.
- [12] D. Cordero-Erasquin, R.J. McCann, M. Michael Schmuckenschläger *Prekopa-Leindler type inequalities on Riemannian manifolds, Jacobi fields, and optimal transport*
- [13] A. Dembo, O. Zeitouni. *Large Deviations Techniques and Applications* Springer. (2nd edition) 1998.
- [14] P. Deift. *Orthogonal Polynomials and Random Matrices: A Riemann-Hilbert Approach* Courant Lect. Notes Math., vol. 3, New York Univ., Courant Inst. Math. Sci., New York, NY; Amer. Math. Soc., Providence, RI, 1999.
- [15] S. Dubuc *Critères de convexité et inégalités intégrales* Annales de l'institut Fourier, tome 27, no 1. p 135-165. 1977.
- [16] R.S. Ellis, K. Haven, B. Turkington *Large Deviation Principles and Complete Equivalence and Nonequivalence Results for Pure and Mixed Ensembles* J. Statist. Phys. 101 no 5-6. p 999-1064. 2000.
- [17] K. Fujita. *On Berman-Gibbs stability of  $\mathbb{Q}$ -Fano varieties* arXiv:1501.00248 [math.AG] 2015.

- [18] C-S Lin, C-L Wang. *Elliptic functions, Green functions and the mean field equations on tori* Ann. of Math. Vol 172. No. 2. 2010.
- [19] J.B. Hough, M. Krishnapur, Y.I: Peres, B. Virag *Determinantal Point Processes and Independence* Probab. Surv. 3. p 206-229. 2006.
- [20] L.V. Kantorovich. *On the Translocation of Masses* C.R. (Dokl) Acad. Sci. URSS 37 (1942), 199-201.
- [21] M. Knott, C.S. Smith. *On the Optimal Mapping of Distributions* Journal of Optimization Theory and Applications: VoL 43, No. 1, May 1984
- [22] R.J. McCann. *Polar factorization of Maps on Riemannian Manifolds* Geom. Funct. Anal. 11, 3 (2001), 589-608.
- [23] J.W. Negele, H. Orland *Quantum Many Particle Systems* Westveiw Press. 1998.
- [24] A. Prekopa *On Logarithmic Concave Measures and functions* Acta Sci. Math., 34. p336-343. 1972.
- [25] A-S Sznitman *Topics in Propagation of Chaos* École d'Été de Probabilités de Saint-Flour XIX-1989. 165-251. Lecture notes in Math. 1464. Springer. 1991.
- [26] C. Villani *Topics in Optimal Transportation* American Mathematical Society. 2003.
- [27] C. Villani *Optimal Transport; Old and New* Springer. 2009.